OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 18:13:48 ; Search time 13687 Seconds
(without alignments)
11466.729 Million cell updates/sec

Title: US-10-045-072-1
Perfect score: 3621
Sequence: 1 tggggcggggttagatcctg.....tgatcgtcgtcgtttcctaa 3621

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		3621	100.0	3621	6	AR123974	AR123974 Sequence
2		3621	100.0	3621	6	AR213173	AR213173 Sequence
3		3621	100.0	3637	1	AF038548	AF038548 Corynebac
4		3621	100.0	332050	1	AP005276	AP005276 Corynebac
5		3621	100.0	349980	6	AX127145	AX127145 Sequence
6		3578	98.8	3728	1	CGPYC	Y09548 Corynebacte
7		3578	98.8	3728	6	A97276	A97276 Sequence 1
8		3578	98.8	3728	6	BD140686	BD140686 Method fo
9		3469	95.8	3657	1	AF503915	AF503915 Corynebac
10		3420	94.4	3420	6	AX120849	AX120849 Sequence
11		3420	94.4	3420	6	BD162966	BD162966 Novel pol
12		3398.8	93.9	3474	6	AX453601	AX453601 Sequence
13		3398.8	93.9	3474	6	AX453603	AX453603 Sequence
14		2282.8	63.0	308750	1	AP005216	AP005216 Corynebac
15		2279.6	63.0	4013	6	BD094153	BD094153 Genes for
16		2279.6	63.0	4851	1	AB083299	AB083299 Corynebac
17		1927.4	53.2	341553	1	BX248355	BX248355 Corynebac
18		1708	47.2	1719	6	AX064993	AX064993 Sequence
19		1395	38.5	1406	6	AX064995	AX064995 Sequence
20	c	1344.6	37.1	291000	1	SCO939105	AL939105 Streptomy
21		1328.2	36.7	4218	1	AF262949	AF262949 Mycobacte
22	c	1294.4	35.7	17992	1	AE007125	AE007125 Mycobacte
23	c	1294.4	35.7	348676	15	BX842581	Bx842581 Mycobacte
24	c	1292.8	35.7	318050	1	BX248344	BX248344 Mycobacte
25		1249.6	34.5	31175	1	MTU00024	U00024 Mycobacteri
26		1249.6	34.5	31176	1	MSGY2	AD000009 Mycobacte
27		1249.6	34.5	32155	6	AR345347	AR345347 Sequence
28		1033	28.5	1083	6	AX064989	AX064989 Sequence
29		1033	28.5	1083	6	AX064991	AX064991 Sequence
30		907	25.0	939	6	AX064985	AX064985 Sequence
31		907	25.0	939	6	AX064987	AX064987 Sequence
32	c	899	24.8	302938	1	BX251412	BX251412 Tropherym
33	c	899	24.8	324227	1	AE016852	AE016852 Tropherym
34	c	741.4	20.5	303862	1	AE017215	AE017215 Geobacter
35		720	19.9	4010	5	AB086371	AB086371 Pagrus ma
36		714	19.7	3994	4	AY185595	AY185595 Bos tauru
37		707.8	19.5	3537	4	AY225510	AY225510 Sus scrof
38	c	705.4	19.5	347786	1	BX640438	BX640438 Bordetell
39		703.2	19.4	3916	8	AF097728	AF097728 Aspergill
40		701.2	19.4	3559	5	AF509529	AF509529 Gallus ga
41	c	699	19.3	349146	1	BX640424	BX640424 Bordetell
42		691.8	19.1	3953	9	HSU04641	U04641 Human pyruv
43		691.8	19.1	4049	9	BC011617	BC011617 Homo sapi
44		690.6	19.1	4067	10	MUSMPYR	L09192 Mus musculu
45		690.2	19.1	3960	9	HSU30891	U30891 Human pyruv

ALIGNMENTS

RESULT 1
AR123974
LOCUS AR123974 3621 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6171833.
ACCESSION AR123974
VERSION AR123974.1 GI:14109335
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3621)
AUTHORS Sinskey,A.J., Lessard,P.A. and Willis,L.B.
TITLE Pyruvate carboxylase from corynebacterium glutamicum
JOURNAL Patent: US 6171833-A 1 09-JAN-2001;
FEATURES Location/Qualifiers

source 1. .3621
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 3621; DB 6; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920

Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980

Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040

Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
```

```
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 2
AR213173

```
LOCUS       AR213173                3621 bp    DNA     linear   PAT 25-SEP-2002
DEFINITION  Sequence 1 from patent US 6403351.
ACCESSION   AR213173
VERSION     AR213173.1  GI:23310240
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3621)
  AUTHORS   Sinskey,A.J., Lessard,P.A. and Willis,L.B.
  TITLE     Pyruvate carboxylase polypeptide from Corynebacterium glutamicum
  JOURNAL   Patent: US 6403351-A 1 11-JUN-2002;
FEATURES             Location/Qualifiers
     source          1. .3621
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match 100.0%; Score 3621; DB 6; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
```

```
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
```

```
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460

Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520

Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580

Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640

Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700

Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760

Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820

Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880

Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940

Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060

Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120

Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180

Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240

Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300

Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360

Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420

Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480

Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540

Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600

Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
             |||||||||||||||||||||
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 3
AF038548

```
LOCUS       AF038548                3637 bp    DNA     linear   BCT 25-AUG-2000
DEFINITION  Corynebacterium glutamicum pyruvate carboxylase (pyc) gene,
            complete cds.
ACCESSION   AF038548
VERSION     AF038548.1  GI:2708716
KEYWORDS    .
SOURCE      Corynebacterium glutamicum
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1  (bases 1 to 3637)
  AUTHORS   Koffas,M.A., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and
            Stephanopoulos,G.
  TITLE     Sequence of the Corynebacterium glutamicum pyruvate carboxylase
            gene
  JOURNAL   Appl. Microbiol. Biotechnol. 50 (3), 346-352 (1998)
  MEDLINE   99019028
   PUBMED   9802220
REFERENCE   2  (bases 1 to 3637)
  AUTHORS   Koffas,M.A.G., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and
            Stephanopoulos,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-DEC-1997) Chemical Engineering, Massachusetts
            Insitute of Technology, 77 Massachusetts Avenue, Cambridge, MA
            02139, USA
FEATURES             Location/Qualifiers
     source          1. .3637
                     /organism="Corynebacterium glutamicum"
                     /mol_type="genomic DNA"
                     /strain="21253"
                     /db_xref="taxon:1718"
     gene            199. .3621
                     /gene="pyc"
     CDS             199. .3621
                     /gene="pyc"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="AAB92588.1"
                     /db_xref="GI:2708717"
                     /translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQL
                     ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQH
                     IEVQILGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI
                     GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEL
                     GLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
                     TGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPL
                     PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA
                     KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
                     YPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDL
                     SDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVH
                     VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLS
                     LEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATALGLADRF
                     ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVI
                     AFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRNSLNR
                     LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
                     PDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVTVA
```

EGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS"

ORIGIN

Query Match 100.0%; Score 3621; DB 1; Length 3637;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920

Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980

Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040

Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
```

```
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160

Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220

Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280

Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340

Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400

Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460

Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520

Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580

Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640

Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700

Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760

Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820

Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880

Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940

Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060

Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120

Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180

Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240

Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300

Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360

Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420

Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480

Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540

Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600

Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 4
AP005276

```
LOCUS       AP005276               332050 bp    DNA     linear   BCT 08-AUG-2002
DEFINITION  Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section
            3/10.
ACCESSION   AP005276 BA000036
VERSION     AP005276.1  GI:21323419
KEYWORDS    .
SOURCE      Corynebacterium glutamicum ATCC 13032
  ORGANISM  Corynebacterium glutamicum ATCC 13032
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1
  AUTHORS   Nakagawa,S.
  TITLE     Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 332050)
  AUTHORS   Nakagawa,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
            Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
            Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,
            Tel:81-44-829-3031, Fax:81-44-813-1651)
COMMENT     This sequence is conducted by collaboration of Kyowa Hakko Kogyo
            Co. Ltd. And Kitasato University.
FEATURES             Location/Qualifiers
     source          1. .332050
                     /organism="Corynebacterium glutamicum ATCC 13032"
                     /mol_type="genomic DNA"
                     /strain="ATCC 13032"
                     /db_xref="taxon:196627"
                     /note="ATCC 13032"
     gene            314. .2518
                     /gene="Cgl0654"
     CDS             314. .2518
                     /gene="Cgl0654"
                     /note="PF02554:Carbon starvation protein CstA"
                     /codon_start=1
                     /transl_table=11
                     /product="Carbon starvation protein, predicted membrane
```

```
                protein"
                /protein_id="BAB98047.1"
                /db_xref="GI:21323420"
                /translation="MKKSPKMSPTARVGLLVFGVIAAVGWGAIAFSRGETINSVWLVL
                AAVGSYIIAFSFYARLIEYKVVKPKDQRATPAEYVNDGKDYVPTDRRVLFGHHFAAIA
                GAGPLVGPVMAAQMGYLPGTLWIILGVIFAGAVQDYLVLWVSTRRRGRSLGQMVRDEM
                GTVGGAAGILATISIMIIIIAVLALIVVNALADSPWGVFSITMTIPIALFMGVYLRYL
                RPGRVTEVSIIGVALLLLAIVAGGWVADTSWGVEWFTWSKTTLALALIGYGIMAAILP
                VWLLLAPRDYLSTFMKIGVIGLLAVGILFARPEVQMPSVTSFALEGNGPVFSGSLFPF
                LFITIACGALSGFHALISSGTTPKLVEKESQMRMLGYGGMLMESFVAMMALITAVILD
                RHLYFSMNAPLALTGGDPATAAEWVNSIGLTGADITPEQLSEAAESVGESTVISRTGG
                APTLAFGMSEILSGFIGGAGMKAFWYHFAIMFEALFILTTVDAGTRVARFMMTDTLGN
                VPGLRRFKDPSWTVGNWISTVFVCALWGAILLMGVTDPLGGINVLFPLFGIANQLLAA
                IALALVLVVVVKKGLYKWAWIPAVPLAWDLIVTMTASWQKIFHSDPAIGYWAQNANFR
                DAKSQGLTEFGAAKSPEAIDAVIRNTMIQGILSILFAVLVLVVVGAAIAVCIKSIRAR
                AAGTPLETTEEPDTESEFFAPTGFLASSRDKEVQAMWDERYPGGAPVSSGGH"
gene            2522. .2725
                /gene="Cgl0655"
CDS             2522. .2725
                /gene="Cgl0655"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical protein"
                /protein_id="BAB98048.1"
                /db_xref="GI:21323421"
                /translation="MMALTHALWKIPRAVWWYLTELMGDTAYSKYVVHLKHHHPDAPI
                PTEREYWRAKYADQDANPGARCC"
gene            complement(3092. .3703)
                /gene="Cgl0656"
CDS             complement(3092. .3703)
                /gene="Cgl0656"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical protein"
                /protein_id="BAB98049.1"
                /db_xref="GI:21323422"
                /translation="MSNVINLVLMQAHSTHQVNLNLIPSSDATDQILATFSSLLCCSQ
                NRRNVISWVGVFRSEERVMEIKLTHRSAVSKRSPFRGITSWQTENRCTPTIGNWPDLC
                HRTGRANWSTQHGSCLHRCIVDDAVDDHFLSVRSNLNWVGCNASDLVSQVLLVRELFR
                RTVGAHFVVNHGWNVTPSQRTNVAHHTNSLKSNVGHCKSCVIL"
gene            3750. .4706
                /gene="Cgl0657"
CDS             3750. .4706
                /gene="Cgl0657"
                /codon_start=1
                /transl_table=11
                /product="Uncharacterized protein involved in propionate
                catabolism"
                /protein_id="BAB98050.1"
                /db_xref="GI:21323423"
                /translation="MWNTIYQAIGQALHTTTATRQSRKGAISSWKAFAPAFAGKMSIE
                AVDRAMRGEGAPSPIWEGEDGVIAWLLSGLDHIYTIPLPAEGEAKRAILDTYTKEHSA
                EYQSQAPIDLARSMGEKLAAQGLDLRDVDSIVLHTSHHTHYVIGTGSNDPQKFDPDAS
                RETLDHSIMYIFAVALEDRAWHHERSYAPERAHRRETIELWNKISTVEDPEWTRRYHS
                VDPAEKAFGARAVITFKDGTVVEDELAVADAHPLGARPFAREQYIQKFRTLAEGVVSE
                KEQDRFLDAAQRTHELEDLSELNIELDADILAKAPVIPEGLF"
gene            4715. .5629
                /gene="Cgl0658"
CDS             4715. .5629
                /gene="Cgl0658"
                /codon_start=1
                /transl_table=11
                /product="PEP phosphonomutase and related enzymes"
                /protein_id="BAB98051.1"
                /db_xref="GI:21323424"
                /translation="MFSSAVAPTERRKALRAALAAPEIARMPGAFSPLAARAIQEAGF
                EGVYVSGAVVAADLALPDIGLTTLTEVAHRSRQIARVTDLPVLVDADTGFGEPMSAAR
                TVSELEDAGVAGCHLEDQVNPKRCGHLDGKEVVGTDIMVRRIAAAVNERRDEQFVICA
                RTDAAGVEGIDSAIERAKAYADAGADMIFTEALYSPADFEKFRAAVDIPLLANMTEFG
                KTELLPAQLLEDIGYNAVIYPVTLLRIAMGQVEQALGDIANTGIQTDWVDRMQHRSRL
                YELLRYNEYNAFDQQVFTYSADSYKPIF"
gene            5658. .6809
                /gene="Cgl0659"
CDS             5658. .6809
                /gene="Cgl0659"
                /EC_number="4.1.3.7"
                /note="PF00285:Citrate synthase"
                /codon_start=1
                /transl_table=11
                /product="Citrate synthase"
                /protein_id="BAB98052.1"
                /db_xref="GI:21323425"
                /translation="MSSATTTDVRKGLYGVIADYTAVSKVMPETNSLTYRGYAVEDLV
                ENCSFEEVFYLLWHGELPTAQQLAEFNERGRSYRSLDAGLISLIHSLPKEAHPMDVMR
                TAVSYMGTKDSEYFTTDSEHIRKVGHTLLAQLPMVLAMDIRRRKGLDIIAPDSSKSVA
                ENLLSMVFGTGPESPASNPADVRDFEKSLILYAEHSFNASTFTARVITSTKSDVYSAI
                TGAIGALKGPLHGGANEFVMHTMLAIDDPNKAAAWINNALDNKNVVMGFGHRVYKRGD
                SRVPSMEKSFRELAARHDGEKWVAMYENMRDAMDARTGIKPNLDFPAGPAYHLLGFPV
                DFFTPLFVIARVAGWTAHIVEQYENNSLIRPLSEYNGEEQREVAPIEKR"
gene            complement(6846. .7040)
                /gene="Cgl0660"
CDS             complement(6846. .7040)
                /gene="Cgl0660"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical protein"
                /protein_id="BAB98053.1"
                /db_xref="GI:21323426"
                /translation="MYFDVSQDVDDELHFSFFELYADQAALDAHRCAPHFYVWREAVA
                KHVLPGSQKNMVGTRMLHHG"
gene            complement(7129. .8172)
                /gene="Cgl0661"
CDS             complement(7129. .8172)
                /gene="Cgl0661"
                /EC_number="1.1.1.82"
                /note="PF02615:Malate/L-lactate dehydrogenase"
                /codon_start=1
                /transl_table=11
                /product="Malate/L-lactate dehydrogenases"
                /protein_id="BAB98054.1"
                /db_xref="GI:21323427"
                /translation="MPEVTVNAQQLTVLCTDILTKTGVPAADAHLVGDSLVQADLWGH
                PSHGVLRLPWYVRRLHSGAMTTHAHVEVLNDLGAVLALDGHNGIGQVLADHARKEAVT
                RAMMFGIGAVSVRNSNHFGTAMYYTRKAAAQGCVSILTTNASPAMAPWGGREKRIGTN
                PWSIAAPFGETATVVDIANTAVARGKIYHARQTNMPIPETWAITSEGAPTTDPAEAIN
                GVVLPMAGHKGYAISFMMDVLSGVLTGSQHSTKVHGPYDPTPPGGAGHLFIALDVAAF
                RDPQDFDDALSDLVGEVKSTPKAQNTEEIFYPGESEDRAHRKNSAHGISLPEKTWMEL
                QELAIENHVVTHR"
gene            complement(8265. .8987)
                /gene="Cgl0662"
CDS             complement(8265. .8987)
                /gene="Cgl0662"
                /note="PF00392:Bacterial regulatory proteins, gntR family"
                /codon_start=1
                /transl_table=11
                /product="Transcriptional regulators"
                /protein_id="BAB98055.1"
                /db_xref="GI:21323428"
                /translation="MVQKDAQASPATRKADQVYTQIRREIEDGTLNPGQRMSEVWLVE
                HTGASRTPVRDALRRLAADELIILEPRQAPMVSPLSLRHIKDLFEFRRIVEVAALEEI
                SVGASKSPRIFGEFSTLAADFRELENSADDADFTADFRRLTSKFDDLVAANTHNQFLG
                RSILSLKPHTTRLRIIAHSDHARLRQSVQEHIEMCEAVASGDLRSAGAACRQHLIHVE
                KSILTALINADSTGSQGIDIRS"
gene            complement(9094. .9798)
                /gene="Cgl0663"
CDS             complement(9094. .9798)
                /gene="Cgl0663"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical membrane protein"
                /protein_id="BAB98056.1"
                /db_xref="GI:21323429"
                /translation="MRFSRVLPALLITTAVSIPTASAATLTADTDKELCIASNTDDSA
                VVTFWNSIEDSVREQRLDELDAQDPGIKAAIESYIAQDDNAPTAAELQVRLDAIESGE
                GLAMLLPDDPTLADPNAEESFKTEYTYDEAKDIISGFSSDPASDVLSQLQQAATTGTR
                TAEIRAEVFADRTDDYNESQTALKEDFQNCIDAIDDARPIPLQYILIGGAIALAVIVL
```

```
                     GIRAWTNSRKQSKHSQ"
     gene            complement(9878. .12094)
                     /gene="Cgl0664"
     CDS             complement(9878. .12094)
                     /gene="Cgl0664"
                     /EC_number="1.1.1.42"
                     /note="TIGR00178:monomer_idh: isocitrate dehydrogenase,
                     NADP-dependent"
                     /codon_start=1
                     /transl_table=11
                     /product="Monomeric isocitrate dehydrogenase (NADP+)"
                     /protein_id="BAB98057.1"
                     /db_xref="GI:21323430"
                     /translation="MAKIIWTRTDEAPLLATYSLKPVVEAFAATAGIEVETRDISLAG
                     RILAQFPERLTEDQKVGNALAELGELAKTPEANIIKLPNISASVPQLKAAIKELQDQG
                     YDIPELPDNATTDEEKDILARYNAVKGSAVNPVLREGNSDRRAPIAVKNFVKKFPHRM

  Query Match             100.0%;  Score 3621;  DB 1;  Length 332050;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3621;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37063 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 37122

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37123 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 37182

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37183 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 37242

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37243 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 37302

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37303 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 37362

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37363 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 37422

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37423 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 37482

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37483 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 37542

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37543 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 37602

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37603 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 37662

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37663 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 37722

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37723 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 37782

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37783 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 37842

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37843 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 37902

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37903 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 37962

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37963 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 38022

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38023 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 38082

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38083 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 38142

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38143 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 38202

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38203 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 38262

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38263 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 38322

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38323 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 38382

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38383 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 38442

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38443 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 38502

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38503 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 38562

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38563 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 38622

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38623 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 38682

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38683 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 38742

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38743 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 38802

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38803 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 38862

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38863 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 38922

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
```

```
Db      38923 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 38982
Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db      38983 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 39042
Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db      39043 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 39102
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db      39103 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 39162
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db      39163 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 39222
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db      39223 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 39282
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db      39283 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 39342
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db      39343 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 39402
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db      39403 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 39462
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db      39463 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 39522
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db      39523 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 39582
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db      39583 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 39642
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db      39643 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 39702
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db      39703 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 39762
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db      39763 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 39822
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db      39823 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 39882
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db      39883 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 39942
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db      39943 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 40002
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
```

```
Db      40003 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 40062
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db      40063 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 40122
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db      40123 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 40182
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db      40183 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 40242
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db      40243 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 40302
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db      40303 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 40362
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db      40363 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 40422
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db      40423 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 40482
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db      40483 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 40542
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db      40543 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 40602
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db      40603 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 40662
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db      40663 TTGATCGTCGTCGTTTCCTAA 40683
```

```
RESULT 5
AX127145
LOCUS       AX127145               349980 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 7061 from Patent EP1108790.
ACCESSION   AX127145 AX114121
VERSION     AX127145.1  GI:14041133
KEYWORDS    .
SOURCE      Corynebacterium glutamicum
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 7061 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Corynebacterium glutamicum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1718"
                     /note="Seq 1 to long (3.309.400) split in 11, seq 7061
                     0.600.001 0.949.980 '"
ORIGIN
```

Query Match 100.0%; Score 3621; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105013 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 105072

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105073 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 105132

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105133 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 105192

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105193 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 105252

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105253 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 105312

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105313 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 105372

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105373 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 105432

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105433 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 105492

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105493 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 105552

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105553 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 105612

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105613 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 105672

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105673 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 105732

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105733 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 105792

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105793 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 105852

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105853 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 105912

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105913 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 105972

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105973 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 106032

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106033 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 106092

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106093 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 106152

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106153 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 106212

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106213 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 106272

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106273 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 106332

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106333 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 106392

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106393 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 106452

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106453 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 106512

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106513 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 106572

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106573 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 106632

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106633 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 106692

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106693 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 106752

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106753 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 106812

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106813 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 106872

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106873 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 106932

Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106933 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 106992

Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106993 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 107052

Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107053 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 107112

Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     107113 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 107172
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db     107173 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 107232
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db     107233 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 107292
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db     107293 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 107352
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db     107353 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 107412
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db     107413 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 107472
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db     107473 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 107532
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db     107533 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 107592
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db     107593 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 107652
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db     107653 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 107712
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db     107713 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 107772
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db     107773 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 107832
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db     107833 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 107892
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db     107893 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 107952
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db     107953 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 108012
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db     108013 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 108072
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db     108073 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 108132
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db     108133 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 108192
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db     108193 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 108252
```

```
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db     108253 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 108312
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db     108313 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 108372
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db     108373 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 108432
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db     108433 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 108492
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db     108493 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 108552
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db     108553 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 108612
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db     108613 TTGATCGTCGTCGTTTCCTAA 108633
```

RESULT 6
CGPYC

```
LOCUS       CGPYC                   3728 bp    DNA     linear   BCT 08-MAY-1998
DEFINITION  Corynebacterium glutamicum pyc gene.
ACCESSION   Y09548
VERSION     Y09548.1  GI:2879822
KEYWORDS    pyc gene; pyruvate carboxylase.
SOURCE      Corynebacterium glutamicum
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1
  AUTHORS   Peters-Wendisch,P.G., Kreutzer,C., Kalinowski,J., Patek,M., Sahm,H.
            and Eikmanns,B.J.
  TITLE     Pyruvate carboxylase from Corynebacterium glutamicum:
            characterization, expression and inactivation of the pyc gene
  JOURNAL   Microbiology (Reading, Engl.) 144 (Pt 4), 915-927 (1998)
  MEDLINE   98240228
   PUBMED   9579065
REFERENCE   2  (bases 1 to 3728)
  AUTHORS   Peters-Wendisch,P.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-1996) P.G. Peters-Wendisch, Institut fuer
            Biotechnologie 1, Forschungszentrum Juelich GmbH, Juelich, D-52425,
            FRG
FEATURES             Location/Qualifiers
     source          1. .3728
                     /organism="Corynebacterium glutamicum"
                     /mol_type="genomic DNA"
                     /strain="ATCC 13032"
                     /db_xref="taxon:1718"
     gene            109. .3587
                     /gene="pyc"
     mRNA            109. .>3587
                     /gene="pyc"
                     /evidence=experimental
     CDS             165. .3587
                     /gene="pyc"
                     /EC_number="6.4.1.1"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="CAA70739.1"
```

/db_xref="GI:2879823"
/db_xref="GOA:O54587"
/db_xref="SPTREMBL:O54587"
/translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
EDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQL
ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEG
QTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQH
IEVQILGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI
GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEL
GLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEIT
AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
TGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPL
PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA
KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
YPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDL
SDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVH
VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLS
LEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATALGLADRF
ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVI
AFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRNSLNR
LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
PDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVTVA
EGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVVS"

ORIGIN

Query Match 98.8%; Score 3578; DB 1; Length 3728;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      44 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 103
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 69

Qy     104 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 163
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 129

Qy     164 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 223
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 189

Qy     224 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 249

Qy     284 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 309

Qy     344 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 369

Qy     404 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 429

Qy     464 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 489

Qy     524 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 549

Qy     584 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 643
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 609

Qy     644 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 703
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 669

Qy     704 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 763
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     670 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 729

Qy     764 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 823
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     730 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 789

Qy     824 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 883
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     790 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 849

Qy     884 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 943
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     850 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 909

Qy     944 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 1003
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     910 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 969

Qy    1004 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     970 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1029

Qy    1064 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1030 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1089

Qy    1124 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1090 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1149

Qy    1184 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1150 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1209

Qy    1244 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1210 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1269

Qy    1304 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1270 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1329

Qy    1364 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1330 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1389

Qy    1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1390 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1449

Qy    1484 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1450 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1509

Qy    1544 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1510 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1569

Qy    1604 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1570 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1629

Qy    1664 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1689

Qy    1724 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1749

Qy    1784 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1843
```

```
Db       1750 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1809
Qy       1844 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1903
Db       1810 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1869
Qy       1904 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1963
Db       1870 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1929
Qy       1964 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 2023
Db       1930 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 1989
Qy       2024 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2083
Db       1990 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2049
Qy       2084 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2143
Db       2050 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2109
Qy       2144 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2203
Db       2110 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2169
Qy       2204 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2263
Db       2170 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2229
Qy       2264 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2323
Db       2230 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2289
Qy       2324 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2383
Db       2290 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2349
Qy       2384 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2443
Db       2350 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2409
Qy       2444 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2503
Db       2410 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2469
Qy       2504 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2563
Db       2470 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2529
Qy       2564 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2623
Db       2530 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2589
Qy       2624 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2683
Db       2590 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2649
Qy       2684 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2743
Db       2650 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2709
Qy       2744 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2803
Db       2710 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2769
Qy       2804 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2863
Db       2770 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2829
Qy       2864 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2923
Db       2830 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2889
Qy       2924 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2983
Db       2890 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2949
Qy       2984 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3043
Db       2950 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3009
Qy       3044 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3103
Db       3010 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3069
Qy       3104 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3163
Db       3070 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3129
Qy       3164 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3223
Db       3130 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3189
Qy       3224 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3283
Db       3190 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3249
Qy       3284 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3343
Db       3250 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3309
Qy       3344 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3403
Db       3310 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3369
Qy       3404 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3463
Db       3370 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3429
Qy       3464 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3523
Db       3430 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3489
Qy       3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3583
Db       3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3549
Qy       3584 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db       3550 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3587
```

```
RESULT 7
A97276
LOCUS       A97276                  3728 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 1 from Patent WO9918228.
ACCESSION   A97276
VERSION     A97276.1  GI:6780661
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 3728)
  AUTHORS   Peters-Wendisch,P. and Eikmanns,B.
  TITLE     METHOD FOR MICROBIAL PRODUCTION OF AMINO ACIDS OF THE ASPARTATE
            AND/OR GLUTAMATE FAMILY AND AGENTS WHICH CAN BE USED IN SAID METHOD
  JOURNAL   Patent: WO 9918228-A 1 15-APR-1999;
            PETERS WENDISCH PETRA (DE); EIKMANNS BERND (DE)
FEATURES             Location/Qualifiers
     source          1. .3728
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN
```

```
  Query Match             98.8%;  Score 3578;  DB 6;  Length 3728;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3578;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         44 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 103
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 69

Qy        104 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 129

Qy        164 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 223
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 189

Qy        224 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 283
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 249

Qy        284 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 343
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 309

Qy        344 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 369

Qy        404 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 463
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 429

Qy        464 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 523
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 489

Qy        524 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 583
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 549

Qy        584 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 643
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 609

Qy        644 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 703
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 669

Qy        704 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 763
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 729

Qy        764 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 823
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 789

Qy        824 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 883
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 849

Qy        884 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 943
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 909

Qy        944 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 1003
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 969

Qy       1004 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1063
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1029

Qy       1064 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1030 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1089

Qy       1124 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1090 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1149

Qy       1184 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1150 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1209

Qy       1244 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1210 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1269

Qy       1304 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1270 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1329

Qy       1364 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1330 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1389

Qy       1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1483
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1390 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1449

Qy       1484 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1543
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1450 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1509

Qy       1544 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1603
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1510 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1569

Qy       1604 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1663
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1570 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1629

Qy       1664 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1723
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1630 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1689

Qy       1724 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1783
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1690 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1749

Qy       1784 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1843
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1750 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1809

Qy       1844 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1903
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1810 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1869

Qy       1904 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1963
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1870 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1929

Qy       1964 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 2023
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1930 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 1989

Qy       2024 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2083
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1990 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2049

Qy       2084 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2050 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2109

Qy       2144 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2203
```

```
Db       2110 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2169
Qy       2204 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2263
Db       2170 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2229
Qy       2264 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2323
Db       2230 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2289
Qy       2324 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2383
Db       2290 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2349
Qy       2384 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2443
Db       2350 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2409
Qy       2444 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2503
Db       2410 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2469
Qy       2504 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2563
Db       2470 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2529
Qy       2564 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2623
Db       2530 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2589
Qy       2624 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2683
Db       2590 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2649
Qy       2684 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2743
Db       2650 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2709
Qy       2744 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2803
Db       2710 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2769
Qy       2804 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2863
Db       2770 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2829
Qy       2864 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2923
Db       2830 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2889
Qy       2924 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2983
Db       2890 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2949
Qy       2984 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3043
Db       2950 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3009
Qy       3044 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3103
Db       3010 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3069
Qy       3104 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3163
Db       3070 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3129
Qy       3164 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3223
Db       3130 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3189
Qy       3224 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3283
Db       3190 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3249
Qy       3284 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3343
Db       3250 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3309
Qy       3344 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3403
Db       3310 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3369
Qy       3404 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3463
Db       3370 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3429
Qy       3464 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3523
Db       3430 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3489
Qy       3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3583
Db       3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3549
Qy       3584 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db       3550 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3587
```

RESULT 8
BD140686

```
LOCUS       BD140686                3728 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Method for microbial production of amino acids of the aspartate
            and/or glutamate family and agents which can be used in said
            method.
ACCESSION   BD140686
VERSION     BD140686.1  GI:23235631
KEYWORDS    JP 2002508921-A/1.
SOURCE      Corynebacterium glutamicum
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1  (bases 1 to 3728)
  AUTHORS   Eikmanns,B., Wendisch,P.P. and Sahm,H.
  TITLE     Method for microbial production of amino acids of the aspartate
            and/or glutamate family and agents which can be used in said method
  JOURNAL   Patent: JP 2002508921-A 1 26-MAR-2002;
            FORSCHUNGSZENTRUM JULICH GMBH
COMMENT     OS   Corynebacterium glutamicum
            PN   JP 2002508921-A/1
            PD   26-MAR-2002
            PF   30-SEP-1998 JP 2000515022
            PR   04-OCT-1997 DE 197 43 894.6,14-JUL-1998 DE 198 31 609.7 PI
            BERND EIKMANNS,PETRA PETERS WENDISCH,HERMANN SAHM PC
            C12N15/09,C12N1/21,C12P13/20//(C12P13/20,C12R1:15),C12N15/00 CC
            pyruvate carboxylase gene
            FH   Key             Location/Qualifiers
            FT   CDS             (165). .(3587).
FEATURES             Location/Qualifiers
     source          1. .3728
                     /organism="Corynebacterium glutamicum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1718"
ORIGIN
```

Query Match 98.8%; Score 3578; DB 6; Length 3728;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          44 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 103
Db          10 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 69
Qy         104 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 163
```

```
Db        70 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 129

Qy       164 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 223
Db       130 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 189

Qy       224 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 283
Db       190 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 249

Qy       284 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 343
Db       250 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 309

Qy       344 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 403
Db       310 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 369

Qy       404 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 463
Db       370 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 429

Qy       464 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 523
Db       430 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 489

Qy       524 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 583
Db       490 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 549

Qy       584 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 643
Db       550 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 609

Qy       644 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 703
Db       610 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 669

Qy       704 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 763
Db       670 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 729

Qy       764 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 823
Db       730 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 789

Qy       824 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 883
Db       790 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 849

Qy       884 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 943
Db       850 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 909

Qy       944 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 1003
Db       910 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 969

Qy      1004 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1063
Db       970 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1029

Qy      1064 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1123
Db       1030 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1089

Qy      1124 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1183
Db       1090 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1149

Qy      1184 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1243
Db       1150 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1209

Qy      1244 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1303
Db       1210 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1269

Qy      1304 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1363
Db       1270 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1329

Qy      1364 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1423
Db       1330 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1389

Qy      1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1483
Db       1390 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1449

Qy      1484 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1543
Db       1450 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1509

Qy      1544 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1603
Db       1510 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1569

Qy      1604 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1663
Db       1570 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1629

Qy      1664 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1723
Db       1630 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1689

Qy      1724 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1783
Db       1690 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1749

Qy      1784 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1843
Db       1750 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1809

Qy      1844 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1903
Db       1810 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1869

Qy      1904 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1963
Db       1870 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1929

Qy      1964 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 2023
Db       1930 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 1989

Qy      2024 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2083
Db       1990 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2049

Qy      2084 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2143
Db       2050 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2109

Qy      2144 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2203
Db       2110 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2169

Qy      2204 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2263
Db       2170 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2229

Qy      2264 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2323
Db       2230 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2289
```

```
Qy       2324 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2290 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2349

Qy       2384 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2350 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2409

Qy       2444 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2410 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2469

Qy       2504 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2470 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2529

Qy       2564 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2530 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2589

Qy       2624 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2590 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2649

Qy       2684 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2650 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2709

Qy       2744 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2710 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2769

Qy       2804 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2770 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2829

Qy       2864 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2830 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2889

Qy       2924 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2890 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2949

Qy       2984 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3043
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2950 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3009

Qy       3044 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3010 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3069

Qy       3104 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3070 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3129

Qy       3164 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3130 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3189

Qy       3224 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3190 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3249

Qy       3284 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3250 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3309

Qy       3344 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3310 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3369

Qy       3404 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3370 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3429

Qy       3464 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3430 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3489

Qy       3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3549

Qy       3584 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||||||||||||||||||||
Db       3550 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3587
```

RESULT 9
AF503915

```
LOCUS       AF503915               3657 bp    DNA     linear   BCT 13-MAY-2002
DEFINITION  Corynebacterium crenatum strain CD945 pyruvate carboxylase (pyc)
            gene, complete cds.
ACCESSION   AF503915
VERSION     AF503915.1  GI:20531756
KEYWORDS    .
SOURCE      Corynebacterium crenatum
  ORGANISM  Corynebacterium crenatum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1  (bases 1 to 3657)
  AUTHORS   Wang,J., Ding,J. and Liu,Y.
  TITLE     Cloning and Expression of Pyruvate Carboxylase Gene in
            Corynebacterium crenatum CD945
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3657)
  AUTHORS   Wang,J., Ding,J. and Liu,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2002) Center of Microbial Biotechnology,
            Institute of Microbiology, Chinese Academy of Sciences, Beijing
            100080, China
FEATURES             Location/Qualifiers
     source          1. .3657
                     /organism="Corynebacterium crenatum"
                     /mol_type="genomic DNA"
                     /strain="CD945"
                     /db_xref="taxon:168810"
     gene            151. .3573
                     /gene="pyc"
     CDS             151. .3573
                     /gene="pyc"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="AAM27458.1"
                     /db_xref="GI:20531757"
                     /translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQL
                     ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQH
                     IEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI
                     GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEL
                     GLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
                     TGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPL
                     PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA
                     KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
                     YPDSVCRAFVKEAATSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDL
                     SDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAATKLVTALRREFDLPVH
                     VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLS
                     LEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATALGLADRF
                     ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVI
                     AFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRNSLNR
                     LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
                     PDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVTVA
```

EGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVIS"

ORIGIN

Query Match 95.8%; Score 3469; DB 1; Length 3657;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3508; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
Qy      49 AAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATG 108
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1 AAGTCGTGCAGGTCAGAGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATG 60

Qy     109 TTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGC 168
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      61 TTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGAAGTGACTGCTATCACCCTTGGT 120

Qy     169 GGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCA 228
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     121 GGTCTCTTGTTGAAAGGAACAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCA 180

Qy     229 GCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCA 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCA 240

Qy     289 GCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTC 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTC 300

Qy     349 CACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCG 408
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     301 CACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGGCTCACCAGTCAAGGCG 360

Qy     409 TACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTAC 468
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     361 TACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTAC 420

Qy     469 CCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGC 528
           |||||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     421 CCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGCGCGGAAAACGGC 480

Qy     529 ATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCG 588
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     481 ATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCG 540

Qy     589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 600

Qy     649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
           |||||||| ||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     601 ATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAAGGCAGTT 660

Qy     709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
           |||||||||||||||||||||||| |||||| |||||||||||||||| ||||||||||||
Db     661 GCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCACCTGATGAGCTCCGCAAATTG 720

Qy     769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
           |||||||||||||||||||||||||||||||| |||||||| || ||||||||||||||||
Db     721 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCGAA 780

Qy     829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
           |||||||||||||||| ||||| |||||||||||||||||||||| ||||||||||||||
Db     781 CGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCACTGGAGAA 840

Qy     889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 900

Qy     949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 960

Qy    1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     961 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1020

Qy    1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1080

Qy    1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1140

Qy    1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1200

Qy    1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1260

Qy    1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1320

Qy    1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1321 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGATTTTGAA 1380

Qy    1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1440

Qy    1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1441 AACATTGGTTTCTTGCGTGCGTTGCTGCGTGAAGAGGACTTCACTTCCAAGCGCATCGCC 1500

Qy    1549 ACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAG 1608
           |||||||| || | |||||||||| ||||||||||||||||||||||| ||||||||||||
Db    1501 ACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCTCCACCTGCGGATGATGAGCAG 1560

Qy    1609 GGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCA 1668
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCA 1620

Qy    1669 AAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGC 1728
           |||||||||||||| || |||||||||||||| ||||||||||||||||||||||||||||
Db    1621 AAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATCAAGGATCTGCCACTGCCACGC 1680

Qy    1729 GGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAG 1788
           |||||||||||||||||||||||||||||| ||||| |||||||| || ||||||||||||
Db    1681 GGTTCCCGTGACCGCCTGAAGCAGCTTGGACCAGCAGCGTTTGCCCGCGATCTCCGTGAG 1740

Qy    1789 CAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCG 1848
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 CAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCG 1800

Qy    1849 ACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCT 1908
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 ACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCT 1860

Qy    1909 GAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTC 1968
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1861 GAGCTTTTGTCCGTGGAGGCCTGGGGCGGTGCGACCTACGATGTGGCGATGCGTTTCCTC 1920

Qy    1969 TTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATT 2028
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1921 TTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTGAACATT 1980

Qy    2029 CAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGC 2088
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    1981 CAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCATACCCAGACTCCGTCTGT 2040

Qy    2089 CGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCG 2148
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    2041 CGCGCGTTTGTTAAGGAAGCTGCCACCTCCGGCGTGGACATCTTCCGCATCTTCGACGCG 2100
```

```
Qy      2149 CTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCG 2208
Db      2101 CTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCG 2160

Qy      2209 GTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTAC 2268
Db      2161 GTCGCTGAAGTGGCTATGGCTTATTCTGGTGATCTTTCCGATCCGAATGAAAAGCTCTAC 2220

Qy      2269 ACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTG 2328
Db      2221 ACCCTGGATTACTACCTGAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATTCTG 2280

Qy      2329 GCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCA 2388
Db      2281 GCTATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCAGCCACCAAGCTGGTCACCGCA 2340

Qy      2389 CTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAG 2448
Db      2341 CTGCGCCGTGAATTTGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAG 2400

Qy      2449 CTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCA 2508
Db      2401 CTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCA 2460

Qy      2509 CCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCAC 2568
Db      2461 CCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCAC 2520

Qy      2569 ACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAA 2628
Db      2521 ACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCATACTGGGAA 2580

Qy      2629 GCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTC 2688
Db      2581 GCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTC 2640

Qy      2689 TACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTG 2748
Db      2641 TACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTG 2700

Qy      2749 GGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTG 2808
Db      2701 GGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCGGCAGTTAATGAGATGCTG 2760

Qy      2809 GGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTC 2868
Db      2761 GGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTC 2820

Qy      2869 GTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCA 2928
Db      2821 GTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCA 2880

Qy      2929 GACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAG 2988
Db      2881 GACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAG 2940

Qy      2989 CCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTT 3048
Db      2941 CCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAAGCACCTTTGACGGAAGTT 3000

Qy      3049 CCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTC 3108
Db      3001 CCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAACAGCCTC 3060

Qy      3109 AACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGC 3168
Db      3061 AACCGCCTGCTGTTCCCGAAGCCAACTGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGC 3120

Qy      3169 AACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACT 3228
Db      3121 AACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACT 3180

Qy      3229 TTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAG 3288
Db      3181 TTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAG 3240

Qy      3289 CCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATG 3348
Db      3241 CCAGACGATAAGGGTATGCGCAATGTTGTGGCTAACGTCAACGGCCAGATCCGCCCAATG 3300

Qy      3349 CGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCC 3408
Db      3301 CGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCC 3360

Qy      3409 AACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGT 3468
Db      3361 AACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACTGTGACTGTTGCTGAAGGT 3420

Qy      3469 GATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3528
Db      3421 GATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3480

Qy      3529 ATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTG 3588
Db      3481 ATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTG 3540

Qy      3589 GAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db      3541 GAAGGTGGCGACTTGATCGTCGTCATTTCCTAA 3573
```

RESULT 10
AX120849
LOCUS AX120849 3420 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 765 from Patent EP1108790.
ACCESSION AX120849
VERSION AX120849.1 GI:14037564
KEYWORDS .
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 765 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1. .3420
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
ORIGIN

Query Match 94.4%; Score 3420; DB 6; Length 3420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       199 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 258
Db         1 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 60

Qy       259 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318
Db        61 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120

Qy       319 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 378
Db       121 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 180

Qy       379 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 438
Db       181 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 240
```

```
Qy      439 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 498
Db      241 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 300

Qy      499 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 558
Db      301 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 360

Qy      559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 618
Db      361 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 420

Qy      619 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 678
Db      421 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 480

Qy      679 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 738
Db      481 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 540

Qy      739 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 798
Db      541 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 600

Qy      799 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 858
Db      601 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 660

Qy      859 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 918
Db      661 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 720

Qy      919 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 978
Db      721 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 780

Qy      979 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 1038
Db      781 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 840

Qy     1039 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 1098
Db      841 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 900

Qy     1099 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 1158
Db      901 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 960

Qy     1159 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1218
Db      961 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1020

Qy     1219 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1278
Db     1021 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1080

Qy     1279 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1338
Db     1081 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1140

Qy     1339 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1398
Db     1141 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1200

Qy     1399 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1458
Db     1201 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1260

Qy     1459 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1518
Db     1261 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1320

Qy     1519 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1578
Db     1321 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1380

Qy     1579 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1638
Db     1381 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1440

Qy     1639 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1698
Db     1441 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1500

Qy     1699 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1758
Db     1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1560

Qy     1759 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1818
Db     1561 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620

Qy     1819 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1878
Db     1621 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680

Qy     1879 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1938
Db     1681 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1740

Qy     1939 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1998
Db     1741 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800

Qy     1999 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 2058
Db     1801 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 1860

Qy     2059 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 2118
Db     1861 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 1920

Qy     2119 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2178
Db     1921 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980

Qy     2179 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2238
Db     1981 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2040

Qy     2239 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2298
Db     2041 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2100

Qy     2299 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2358
Db     2101 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2160

Qy     2359 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2418
Db     2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2220

Qy     2419 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2478
Db     2221 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2280

Qy     2479 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2538
Db     2281 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2340

Qy     2539 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2598
Db     2341 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2400

Qy     2599 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2658
```

```
Db       2401 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2460
Qy       2659 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2718
Db       2461 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2520
Qy       2719 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2778
Db       2521 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2580
Qy       2779 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2838
Db       2581 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2640
Qy       2839 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2898
Db       2641 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2700
Qy       2899 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2958
Db       2701 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2760
Qy       2959 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 3018
Db       2761 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 2820
Qy       3019 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 3078
Db       2821 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 2880
Qy       3079 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 3138
Db       2881 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 2940
Qy       3139 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3198
Db       2941 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3000
Qy       3199 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3258
Db       3001 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3060
Qy       3259 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3318
Db       3061 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3120
Qy       3319 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3378
Db       3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3180
Qy       3379 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3438
Db       3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3240
Qy       3439 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3498
Db       3241 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3300
Qy       3499 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3558
Db       3301 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360
Qy       3559 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3618
Db       3361 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3420

RESULT 11
BD162966
LOCUS       BD162966               3420 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Novel polynucleotide.
ACCESSION   BD162966
VERSION     BD162966.1  GI:27868728
KEYWORDS    JP 2002191370-A/765.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 3420)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotide
  JOURNAL   Patent: JP 2002191370-A 765 09-JUL-2002;
            KYOWA HAKKO KOGYO CO LTD
COMMENT     OS   Corynebacterium glutamicum
            PN   JP 2002191370-A/765
            PD   09-JUL-2002
            PF   15-DEC-2000 JP 2000405096
            PI   SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
            PI   KEIKO OCHIAI,
            PI   HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
            PI   OZAKI
            PC   C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00, PC
              C12N1/15,
            PC   C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
            PC   04,C12P13/08,
            PC   C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53, PC
             G01N33/566,
            PC   G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
            PC   (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
            PC   C12N15/00,
            PC   C12N5/00,C12N15/00
            CC   Novel polynucleotide
            FH   Key             Location/Qualifiers
            FT   source          1. .3420
            FT                   /organism='Corynebacterium glutamicum'.
FEATURES             Location/Qualifiers
     source          1. .3420
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             94.4%;  Score 3420;  DB 6;  Length 3420;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3420;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        199 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 258
Db          1 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 60
Qy        259 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318
Db         61 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120
Qy        319 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 378
Db        121 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 180
Qy        379 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 438
Db        181 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 240
Qy        439 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 498
Db        241 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 300
Qy        499 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 558
Db        301 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 360
Qy        559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 618
Db        361 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 420
Qy        619 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 678
```

```
Db        421 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 480
Qy        679 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 738
Db        481 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 540
Qy        739 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 798
Db        541 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 600
Qy        799 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 858
Db        601 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 660
Qy        859 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 918
Db        661 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 720
Qy        919 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 978
Db        721 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 780
Qy        979 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 1038
Db        781 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 840
Qy       1039 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 1098
Db        841 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 900
Qy       1099 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 1158
Db        901 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 960
Qy       1159 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1218
Db        961 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1020
Qy       1219 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1278
Db       1021 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1080
Qy       1279 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1338
Db       1081 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1140
Qy       1339 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1398
Db       1141 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1200
Qy       1399 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1458
Db       1201 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1260
Qy       1459 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1518
Db       1261 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1320
Qy       1519 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1578
Db       1321 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1380
Qy       1579 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1638
Db       1381 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1440
Qy       1639 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1698
Db       1441 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1500
Qy       1699 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1758
Db       1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1560
```

```
Qy       1759 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1818
Db       1561 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620
Qy       1819 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1878
Db       1621 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680
Qy       1879 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1938
Db       1681 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1740
Qy       1939 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1998
Db       1741 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800
Qy       1999 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 2058
Db       1801 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 1860
Qy       2059 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 2118
Db       1861 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 1920
Qy       2119 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2178
Db       1921 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980
Qy       2179 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2238
Db       1981 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2040
Qy       2239 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2298
Db       2041 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2100
Qy       2299 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2358
Db       2101 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2160
Qy       2359 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2418
Db       2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2220
Qy       2419 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2478
Db       2221 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2280
Qy       2479 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2538
Db       2281 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2340
Qy       2539 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2598
Db       2341 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2400
Qy       2599 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2658
Db       2401 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2460
Qy       2659 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2718
Db       2461 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2520
Qy       2719 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2778
Db       2521 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2580
Qy       2779 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2838
Db       2581 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2640
```

```
Qy       2839 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2700

Qy       2899 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2760

Qy       2959 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 3018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 2820

Qy       3019 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 3078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 2880

Qy       3079 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 3138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 2940

Qy       3139 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3000

Qy       3199 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3060

Qy       3259 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3120

Qy       3319 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3180

Qy       3379 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3240

Qy       3439 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3300

Qy       3499 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360

Qy       3559 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3420
```

```
RESULT 12
AX453601
LOCUS       AX453601                3474 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 1 from Patent WO0231158.
ACCESSION   AX453601
VERSION     AX453601.1  GI:21712841
KEYWORDS    .
SOURCE      Corynebacterium glutamicum
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1
  AUTHORS   Hanke,P.D.
  TITLE     Feedback-resistant pyruvate carboxylase gene from corynebacterium
  JOURNAL   Patent: WO 0231158-A 1 18-APR-2002;
            ARCHER-DANIELS-MIDLAND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .3474
                     /organism="Corynebacterium glutamicum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1718"
     CDS             1. .3474
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD38136.1"
                     /db_xref="GI:21712842"
                     /db_xref="REMTREMBL:CAD38136"
                     /translation="MTAITLGGLLLKGIITLVSTHTSSTLPAFKKILVANRGEIAVRA
                     FRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVK
                     ADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLA
                     ESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAA
                     FGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDP
                     ELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVD
                     LVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPG
                     GAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIG
                     FLRALLREEDFTSKRIATGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPK
                     DVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLL
                     ATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPN
                     VNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVL
                     ETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA
                     VTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSL
                     SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQ
                     LSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDP
                     ADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQ
                     AHLDADDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIR
                     LPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSN
                     KGHVAAPFAGVVTVTVAEGDEVKAGDAVAIIEAMKMEATITASVDGKIERVVVPAATK
                     VEGGDLIVVVS"
ORIGIN

  Query Match             93.9%;  Score 3398.8;  DB 6;  Length 3474;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 3427;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;
```

```
Qy        148 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy        208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120

Qy        268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180

Qy        328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240

Qy        388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300

Qy        448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
              |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy        508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420

Qy        568 ACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 480

Qy        628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540

Qy        688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
              |||||||||| |||||||||||||||||||||||||||||||||||| |||||| |||||
```

```
Db        541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCA 600
Qy        748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
Db        601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660
Qy        808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
Db        661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720
Qy        868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
Db        721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
Qy        928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
Db        781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840
Qy        988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
Db        841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
Qy       1048 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107
Db        901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC 960
Qy       1108 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167
Db        961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020
Qy       1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227
Db       1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
Qy       1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287
Db       1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140
Qy       1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347
Db       1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200
Qy       1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407
Db       1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
Qy       1408 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467
Db       1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320
Qy       1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527
Db       1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1380
Qy       1528 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1587
Db       1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT 1440
Qy       1588 CCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647
Db       1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
Qy       1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
Db       1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
Qy       1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
Db       1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1620
Qy       1768 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827
Db       1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
```

```
Qy       1828 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887
Db       1681 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
Qy       1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1947
Db       1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1800
Qy       1948 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 2007
Db       1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860
Qy       2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067
Db       1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920
Qy       2068 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127
Db       1921 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
Qy       2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2187
Db       1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
Qy       2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247
Db       2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
Qy       2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307
Db       2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
Qy       2308 AAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2367
Db       2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220
Qy       2368 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427
Db       2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
Qy       2428 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487
Db       2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
Qy       2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2547
Db       2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2400
Qy       2548 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607
Db       2401 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Qy       2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2667
Db       2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
Qy       2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727
Db       2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
Qy       2728 CGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC 2787
Db       2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy       2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847
Db       2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
Qy       2848 GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2907
Db       2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760
```

```
Qy       2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2820

Qy       2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 3027
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 2880

Qy       3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 3087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 2940

Qy       3088 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3147
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000

Qy       3148 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC 3207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060

Qy       3208 CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3267
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120

Qy       3268 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180

Qy       3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240

Qy       3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300

Qy       3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3507
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360

Qy       3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 3567
             |||||||||||||||||||||||||||||||||||||||||| || || |||||||||
Db       3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420

Qy       3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3474
```

RESULT 13
AX453603
LOCUS AX453603 3474 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 3 from Patent WO0231158.
ACCESSION AX453603
VERSION AX453603.1 GI:21712843
KEYWORDS .
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Hanke,P.D.
TITLE Feedback-resistant pyruvate carboxylase gene from corynebacterium
JOURNAL Patent: WO 0231158-A 3 18-APR-2002;
ARCHER-DANIELS-MIDLAND COMPANY (US)
FEATURES Location/Qualifiers
source 1. .3474
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
ORIGIN

Query Match 93.9%; Score 3398.8; DB 6; Length 3474;

Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy        148 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy        208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120

Qy        268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180

Qy        328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240

Qy        388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300

Qy        448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
              |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy        508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420

Qy        568 ACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 480

Qy        628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540

Qy        688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
              |||||||||| |||||||||||||||||||||||||||||||||| |||||| |||||
Db        541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCA 600

Qy        748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
              |||||||||| |||||||| |||||||||||||||||||||||||||||| |||||||
Db        601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660

Qy        808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
              || ||  |||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db        661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720

Qy        868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780

Qy        928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840

Qy        988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900

Qy       1048 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC 960

Qy       1108 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020

Qy       1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
Qy       1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287
Db       1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140
Qy       1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347
Db       1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200
Qy       1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407
Db       1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
Qy       1408 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467
Db       1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320
Qy       1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527
Db       1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1380
Qy       1528 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1587
Db       1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT 1440
Qy       1588 CCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647
Db       1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
Qy       1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
Db       1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
Qy       1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
Db       1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1620
Qy       1768 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827
Db       1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
Qy       1828 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887
Db       1681 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
Qy       1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1947
Db       1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1800
Qy       1948 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 2007
Db       1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860
Qy       2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067
Db       1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920
Qy       2068 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127
Db       1921 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
Qy       2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2187
Db       1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
Qy       2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247
Db       2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
Qy       2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307
Db       2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
```

```
Qy       2308 AAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2367
Db       2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220
Qy       2368 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427
Db       2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
Qy       2428 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487
Db       2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
Qy       2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2547
Db       2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2400
Qy       2548 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607
Db       2401 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Qy       2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2667
Db       2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
Qy       2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727
Db       2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
Qy       2728 CGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC 2787
Db       2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy       2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847
Db       2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
Qy       2848 GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2907
Db       2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760
Qy       2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2967
Db       2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2820
Qy       2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 3027
Db       2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 2880
Qy       3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 3087
Db       2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 2940
Qy       3088 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3147
Db       2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000
Qy       3148 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC 3207
Db       3001 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060
Qy       3208 CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3267
Db       3061 CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
Qy       3268 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3327
Db       3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180
Qy       3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387
Db       3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240
```

```
Qy       3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300

Qy       3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3507
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360

Qy       3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 3567
             |||||||||||||||||||||||||||||||||||||||||||| || || |||||||||
Db       3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420

Qy       3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3474
```

RESULT 14
AP005216
LOCUS AP005216 308750 bp DNA linear BCT 10-JUL-2003
DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 3/11.
ACCESSION AP005216 BA000035
VERSION AP005216.1 GI:23492422
KEYWORDS .
SOURCE Corynebacterium efficiens YS-314
ORGANISM Corynebacterium efficiens YS-314
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nishio,Y., Nakamura,Y., Kawarabayasi,Y., Usuda,Y., Kimura,E., Sugimoto,S., Matsui,K., Yamagishi,A., Kikuchi,H., Ikeo,K. and Gojobori,T.
TITLE Comparative Complete Genome Sequence Analysis of the Amino Acid Replacements Responsible for the Thermostability of Corynebacterium efficiens
JOURNAL Genome Res. 13 (7), 1572-1579 (2003)
MEDLINE 22723752
PUBMED 12840036
REFERENCE 2 (bases 1 to 308750)
AUTHORS Kawarabayasi,Y., Yamazaki,J., Hino,Y., Kikuchi,H. and Director-General of Biotechnology Center.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bio@nite.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
COMMENT Kawarabayasi, Y is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.
FEATURES Location/Qualifiers
source 1. .308750
/organism="Corynebacterium efficiens YS-314"
/mol_type="genomic DNA"
/strain="YS-314"
/db_xref="taxon:196164"
CDS 155. .598
/note="CE0586, similar to AE007159-8|AAK47889.1| percent identity: 63 in 145 aa"
/codon_start=1
/transl_table=11
/product="putative 50S ribosomal protein L13"
/protein_id="BAC17396.1"
/db_xref="GI:23492423"
/translation="MSTYHPKSGDITRKWYVIDATDVVLGRLATHAADLLRGKGKPLF APNVDCGDHVIIINADKVAITSNKRDREMRYRHSGYPGGLKSMTLGRSLELHPERVLE ESITGMMPHNKLTAASVKKLHVFAGSEHPYAAQKPETYEIKKVAQ"
CDS 598. .1146
/note="CE0587, similar to AL583918-48|CAC29873.1| percent identity: 59 in 159 aa"
/codon_start=1
/transl_table=11
/product="putative 30S ribosomal protein S9"
/protein_id="BAC17397.1"
/db_xref="GI:23492424"
/translation="MSEPIQNENVESNVADAADIAAATAATEEFTNTIGDAISTSTEE ETFEAAPAVLDGPIWTVGRRKRAIVRVRMVAGSGEITCNGRTLEDYFPNKLHQQLIKA PLVLLGREGQFDIHANLGGGGPTGQAGAFRLAIARALNAYNPAERPTLKKAGFLTRDA RAVERKKAGLHKARRAPQYSKR"
CDS 1282. .2670
/note="CE0588, similar to AL031317-14|CAA20392.1| percent identity: 62 in 448 aa"
/codon_start=1
/transl_table=11
/product="putative phospho-sugar mutase"
/protein_id="BAC17398.1"
/db_xref="GI:23492425"
/translation="MWCRAGGRWTCIMLGMTRLFGTDGVRGLANKTLTAPLALKLGAA AAHVLTAGTRPHGRRPVAIVGRDPRVSGEMLAAALAAGMASRGVDVLRVGVIPTPGVA FLTDDYGADMGVMISASHNPMPDNGIKFFSAGGHKLPDEVEDEIERVMDDLPEEGPTG HGIGRVIEEAPDARGRYLQHLADAVPTDLSGITVVVDAANGAASVIAPQAYEAAGAKV IAIHNTPNAYNINEKCGSTHMDQIQAAVLEHGADLGLAHDGDADRCLAVDSDGNIVDG DQIMAILAIAMKENSELRKNTLVATVMSNLGLKLAMEKADIQLRTTKVGDRYVLEELN AGGFALGGEQSGHIVLPDHGTTGDGTLTGLSLMARMAATGKPLSELAAAMTVLPQVLI NVPVADKSSIMKSANVQAAVAAAEEELGGTGRVLLRPSGTEELFRVMVEAAEQEQARR VAGRLAAVVAEA"
CDS 2742. .3071
/note="CE0589"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17399.1"
/db_xref="GI:23492426"
/translation="MLRGIVMSGLNLNVEGALTAARTMLGEATHQRQRHRVAPPAFPV SAAGQGFSTWGAQVATMLEELHRGVDKRLDAVEATADAAAREVGVFRDADAGIGRGLG QVDGEGR"
CDS 3068. .4462
/note="CE0590"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17400.1"
/db_xref="GI:23492427"
/translation="MSASISSYGSAVSQLQAASRGTYPGPSIPERVLKKAYETITGLD TDQLRSGGRSAAGDSETGGLLGLLKPLGRMFGGVAGGMIINELVDRVSDWVDNRNEAE EVAEAAERAADAIDTTIRESDDGVEAILGQLAGIIAEISNHLSTLDPVEHPEAFIATV QAGAEIINDAAAMVLGLCGDRDRAIESCYGALIEHGRRVCEAPAPGVAPAVGGGGGGG AGTIPESPVTPPVTTSGATPGTAPASVDPPTADSPEAGTATGRTEAEDCDPEPGQDEK PVESPAEPPDEDETGVEPPSPEDDDWEHTLTGAISGIGIGLLVVGVGALVDFLGEVAQ TMGGDAAQMPPVPSPGAAPLAATPTVEPVPAPAPESAPEPNPEPTPKPTPASELHLVA EPPPKPQPGMVQAAATAAVQEPTVGAPVQVDYQQAATPAAPVAPPTAAPAPGNSSGDA AGDAPGRVRRVGGW"
CDS 4456. .4716
/note="CE0591"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17401.1"
/db_xref="GI:23492428"
/translation="MVAVEYEVFVAQFRQRTAARMAEFEKALAKAQGEVAKTSPGEPG RRQAAMSPPQVTPASPVARTPQQPGRGRPRGRGPVQSILRRA"
CDS complement(4743. .5663)
/note="CE0592, similar to AL583918-51|CAC29878.1| percent identity: 44 in 254 aa"
/codon_start=1
/transl_table=11

```
                /product="conserved hypothetical protein"
                /protein_id="BAC17402.1"
                /db_xref="GI:23492429"
                /translation="MAHPGRIILPSGVTLSTILDRVAENLKKHLSKLSKRGPHRVMVG
                DLDYAGLPGKIYTPAEGNGVPGVAFGHDWMKPIKHYHQTLRHLASWGIAVAAPNTETG
                FLPDHRGFAADLDSALQILAGVKLGAGKVTVNPGGLGLVGHGMGGGVAVLAAANRKAV
                RAVGALYPASTSPSCVDAAPAVKAPGLIIGTADLGMFDAGDPAKVAARWGGDVVYREL
                DNGNQQGFSEDTMFKLLVGVGRPQTAGQELARGLLTGFLLHQLAGERSYKAFSDPEAE
                AKKVTSYWGIDLEEKAFPQDESPLPFLNSN"
gene            5737. .6840
                /gene="alr"
CDS             5737. .6840
                /gene="alr"
                /EC_number="5.1.1.1"
                /note="CE0593, similar to AX063847-1|CAC25165.1| percent
                identity: 68 in 362 aa"
                /codon_start=1
                /transl_table=11
                /product="alanine racemase"
                /protein_id="BAC17403.1"
                /db_xref="GI:23492430"
                /translation="MNLLTTRIDLDAIAHNTRLLKNRVGAAKLMAVVKADGYNHGMDR
                VAPVMAANGADAFGVATIAEALALRATGITTPILCWIWSPEQDWAAAVEAGIDLAVIS
                PRHARVLVDAPPTSRAIRVSVKVDTALHRSGVDEQDWDAVFTLLRDCGHIEVTGLFTH
                LSCADEPGNPETDHQAETFRRAIDRARALGLEVPENHLCNSPATLTRPDLHMDMVRPG
                VALYGLEPIPGLDHGLRPAMTWAGAITVVKPIRKGEGTSYGLTWRAEADGFVAVVPAG
                YADGVPRAAQDHLRVHVNGHDYPQVGRVCMDQFVIFLGDNPHGVTAGDEAVIFGGTGM
                SATELADALATINYEVICRPTGRTVRDYTGEGK"
CDS             6837. .7340
                /note="CE0594, similar to AL583918-56|CAC29885.1| percent
                identity: 50 in 156 aa"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAC17404.1"
                /db_xref="GI:23492431"
                /translation="MKPDFPASGTRRLDTAAQTQAFGEELGRALEAGDVVVLDGPLGA
                GKTTFTQGLARGMQVRGRVTSPTFVIAREHRSEVGGPTLIHLDAYRLLGEDPDDVDSD
                PIGALDSLDLDTDLDTAVVVAEWGGGLVEQITESYLLISIDRETAVTEDPESEARIFT
                WQWRSRH"
CDS             7540. .9132
                /note="CE0595, similar to AX066613-1|CAC26534.1| percent
                identity: 41 in 530 aa"
                /codon_start=1
                /transl_table=11
                /product="putative transport protein"
                /protein_id="BAC17405.1"
                /db_xref="GI:23492432"
                /translation="MLDFLAANPLIALAVILAVGLAIGRISLFGVSLGAAAVLIVALV
                VSTLNPDIQIPAFVFQLGLAMFVYVIGISAGPAFFREFRSRGWKLTLFMITLLVSLTA
                LAWVLIRAFGLGAGAGAGMFAGSLTSTPGMAAVVALMDPAQAGDPVIGYSLAYPGAVL
                GSILVAAVGAKLLKVNHTEDAREEGLVTEPLVWKGVRIGEGISGTIGDLPRLSGQQII
                ATRIVEDPHEHRLADPTLPIKPGMELVINGTVNAVDRAIAALGGECDTKIEDTELVYS
                RFTVSNPDIVGRTVAELDPVANGFMIARIRQGDTEIVPHRDTVLNYSDRVRVVAAPGR
                MGEVRRFLGDSEKALGDVNLLPFAIGLSLGLLLGAIPVPLPGDTTMYLGFGGGPIVAG
                LILGALNRTGPITWQLPFHANRTISTLGLALFLAGVGTSAGAGFRQALTDPQSFVYMG
                VGFAITVTSALVCAVVGMWLLKLKWDESMGVAAGATTNPAIISYLNDQTGTDLANRGY
                ATVYPTAMIGKILACQVLFLLL"
CDS             9183. .10529
                /note="CE0596, similar to AX065751-1|CAC26115.1| percent
                identity: 79 in 430 aa"
                /codon_start=1
                /transl_table=11
                /product="putative prolyl aminopeptidase"
                /protein_id="BAC17406.1"
                /db_xref="GI:23492433"
                /translation="MARPGGPTGWPDWGLTGGGPVSTRGRNLDGMTELSTTHLGELRL
                DELTLTVPLTTDVADERTIDIFARIATRPGGETLPYLVFLQGGGPGNEAPRPSLEPLNP
                SWLKAALEHYRVVMLDQRGTGLSTPVSDRILEEHTTAEIVEYLSHLRADGIVRDCEAV
                REYLGVRQWNVLGQSFGGFTTLHYLSTHPDSLDDVFITGGLSAIDRPAEDVYANCYER
                MRYNSEQYYRRFPEDRVGIQRLVERARAGEIILPTGEVVSESRLRSLGHLLGSNDGWL
                DLHHLLERDPRSNAFLHDLADLLPFGARNPLYYVLHESSYADGVVTDWAAQRVYPEAF
                REDPTLLTGEHVFPEWADTVPALRPWKDVALALAQQEWPKLYDAAALETSGATGAAAV
```

```
Query Match             63.0%;  Score 2282.8;  DB 1;  Length 308750;
Best Local Similarity   77.7%;  Pred. No. 0;
Matches 2810;  Conservative    0;  Mismatches  797;  Indels    9;  Gaps    4;

Qy          7 GGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGGTCAGGG 66
              |||||      | |||||  || |||      |    |   ||||  |   ||||||||
Db     128583 GGGGTCTCGGGCGGGGGGCATTCTTTTCACGGCAAGGTGGTGAAATTCCGCAGGTCACTC 128642

Qy         67 GAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGGAATCGG 126
                | | | | | | |   ||||||  ||| | ||   || ||   | | ||
Db     128643 CCCGGCCGGCGGTAGAGAACGGAGCGAAAACGGAAAGCAATACGTGGTTTTCCGGACTGG 128702

Qy        127 GGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTT--GGCGGTCTCTTGTTGAAAG 184
              ||||||||  |  || | |||||||| |||||||    ||| |   || || |||||
Db     128703 CCGTTACGATGTTCTGAAG-AGTGACTGCCATCACCCAACAGGCTGGTCCTCGTCGAAAG 128761

Qy        185 GAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCT 244
              |||  ||  |||  |  ||  |  ||    |||  |  ||  |||||  ||  || ||||||||||||
Db     128762 GAACAAAAACTGTGGTTACAACA---ACACCCTCCACGCTGCCGGCGTTCAAAAAGATCC 128818

Qy        245 TGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTG 304
              ||||  ||  |||||  ||  ||||||||||  ||  ||  |||||  ||  ||   |||  |||||  |
Db     128819 TGGTGGCCAACCGAGGTGAAATCGCGGTGCGAGCATTCCGCGCCGCCTACGAGACCGGGG 128878

Qy        305 CAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTT 364
              | ||  ||  ||  ||  ||  ||||||||  ||  ||  ||  ||  ||  ||||||||||||  ||  ||  |
Db     128879 CCGCAACCGTGGCCATCTACCCCCGGGAGGACCGTGGCTCCTTCCACCGCTCCTTCGCCT 128938

Qy        365 CTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATG 424
              | ||  ||  ||   |  ||  ||  ||||  ||||||||  ||  ||||||||||||  ||  ||||  |
Db     128939 CCGAGGCGGTGAGGATCGGAACCGAGGGCTCACCCGTCAAGGCGTACCTCGATATTGATG 128998

Qy        425 AAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCC 484
              | ||  |||    ||  ||  ||  ||  ||  |||||  ||  ||   |  ||||||||  ||  ||  ||||
Db     128999 AGATCATCAACGCCGCCAAGAAGGTGAAAGCGGACGCGGTCTACCCGGGGTATGGTTTCC 129058

Qy        485 TGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCC 544
              | ||  ||||||||||||||  ||  ||  ||  |||||  ||||||||||||  ||  ||  |
Db     129059 TTTCGGAAAATGCCCAGCTCGCGCGTGAATGCGCGGAGAACGGCATTACCTTCATCGGTC 129118

Qy        545 CAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGA 604
              | |||||  |||||  ||  ||  |||||  ||  ||  |||||      ||  ||  ||||||||||||
Db     129119 CCACCCCGGAGGTGCTCGACCTCACGGGCGACAAGTCCAAGGCTGTGTCCGCCGCGAAGA 129178

Qy        605 AGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTA 664
              ||||  ||  |||||  ||  ||||||||||||||||  ||||    ||||||||||||||||  |
Db     129179 AGGCCGGGCTGCCGGTGCTGGCGGAATCCACCCCCAGCACCGACATCGATGAGATCGTCA 129238

Qy        665 AAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGAC 724
              | ||  ||  ||  ||  |||||  |||||  |||||  ||  |||||  ||  ||  ||||||||||||  |
Db     129239 AGAGTGCCGAGGGGCAGACCTACCCGATCTTCGTCAAGGCCGTCGCAGGTGGTGGCGGGC 129298

Qy        725 GCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTC 784
              | ||||||||  ||  ||  |    ||  ||  ||  ||  ||  |  |  ||  |  ||  ||  ||  |
Db     129299 GTGGTATGCGGTTCGTCGAGAAGCCCGAGGACCTGCGTGAGCTGGCCAGGGAGGCCTCCC 129358

Qy        785 GTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACC 844
              | ||  ||  ||  ||  ||||||||  ||  ||  |  ||  ||  ||||||||  ||  |||||  ||  |
Db     129359 GCGAGGCGGAGGCCGCTTTCGGTGACGGATCCGTCTACGTCGAACGGGCCGTGATCAAAC 129418

Qy        845 CTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATG 904
              | |||||  ||  ||  ||||||||||  ||  ||||||||  ||  ||  ||  |  |||||  ||||
Db     129419 CCCAGCACATCGAGGTGCAGATCCTCGGTGATCACACCGGCGATGTCATCCACCTGTATG 129478

Qy        905 AACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGC 964
              ||||  |||||  ||  ||||||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||||
Db     129479 AACGCGACTGTTCCCTGCAGCGCCGCCACCAGAAGGTCGTGGAGATCGCACCTGCCCAGC 129538

Qy        965 ATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCA 1024
              |  |  ||  ||  ||  |||||  ||  |||||  |||||  |||||  ||  |||||||||    ||||
Db     129539 ACCTCGACCCGGAGCTGCGCGACCGCATCTGTGCCGATGCCGTGAAGTTCTGCAAATCCA 129598
```

```
Qy       1025 TTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCT 1084
Db     129599 TCGGATACCAGGGCGCCGGCACCGTGGAGTTCCTCGTCGACGAGGCGGGCAACCACGTCT 129658

Qy       1085 TCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGG 1144
Db     129659 TCATTGAGATGAACCCCCGCATCCAGGTGGAACACACCGTGACCGAGGAGGTCACCTCCG 129718

Qy       1145 TGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTC 1204
Db     129719 TCGACCTGGTCAAGGCGCAGATGCACCTGGCCGCCGGTGCCACCCTGAAGGAACTGGGCC 129778

Qy       1205 TGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAG 1264
Db     129779 TGACCCAGGACAAGATCACCACCCACGGTGCCGCCCTGCAGTGCCGCATCACCACGGAGG 129838

Qy       1265 ATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCG 1324
Db     129839 ACCCGTCCAACAACTTCCGGCCCGACACCGGTGTGATCACCGCCTACCGCTCCCCGGGTG 129898

Qy       1325 GAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTG 1384
Db     129899 GTGCGGGTGTGCGTCTCGACGGCGCAGCCCAGCTCGGCGGCGAGATCACCGCACATTTCG 129958

Qy       1385 ACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTG 1444
Db     129959 ATTCCATGCTGGTCAAGATGACCTGCCGCGGTTCCGATTTCGAGACCGCCGTGTCCCGAG 130018

Qy       1445 CACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGC 1504
Db     130019 CCCAGCGCGCCCTGGCGGAGTTCAACGTCTCCGGCGTGGCCACCAACATCGGCTTCCTGC 130078

Qy       1505 GTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCG 1564
Db     130079 GTGCGCTGCTGCGCGAGGAAGACTTCACCAAGAGGCGCATCGACACCGGCTTCATCGGCT 130138

Qy       1565 ATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATT 1624
Db     130139 CCCACCAGCACCTGCTCCAGGCCCCACCGGCCGACGATGAGCAGGGGCGGATCCTGGAAT 130198

Qy       1625 ACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTC 1684
Db     130199 ACCTGGCGGATGTCACCGTGAACAAACCCCACGGTGAACGCCC---CGAGACAGCCCGTC 130255

Qy       1685 CTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCC 1744
Db     130256 CGATAGAGAAGCTGCCCGAGGTGGAGAACATCCCGCTGCCACGCGGCTCCCGCGACCGCC 130315

Qy       1745 TGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAG 1804
Db     130316 TGAAGCAGCTCGGCCCGGAGGGTTTCGCCCGCGATCTGCGCGAACAGGATGCCCTGGCCG 130375

Qy       1805 TTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCAT 1864
Db     130376 TCACCGACACCACCTTCCGCGATGCCCACCAGTCCCTCCTGGCCACCCGCGTGCGCTCCT 130435

Qy       1865 TCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGG 1924
Db     130436 TCGCGCTGACCCCGGCGGCGCGCGCCGTCGCAAAGCTCACCCCCGAGCTGCTGTCGGTGG 130495

Qy       1925 AGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGG 1984
Db     130496 AGGCCTGGGGCGGTGCCACCTACGACGTGGCCATGCGCTTCCTCTTCGAGGATCCGTGGG 130555

Qy       1985 ACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCG 2044
Db     130556 CACGCCTGGATGAGCTGCGTGAGGCGATGCCGAATGTGAACATCCAGATGCTGCTGCGTG 130615

Qy       2045 GCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGG 2104
Db     130616 GTCGCAACACCGTCGGGTACACCCCGTACCCCGATTCGGTGTGCCGCGCGTTTGTGCAGG 130675

Qy       2105 AAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCC 2164
Db     130676 AGGCCGCCAAGTCCGGTGTGGACATCTTCCGCATCTTCGACGCGCTCAACGACATCTCCC 130735

Qy       2165 AGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTA 2224
Db     130736 AGATGCGCCCGGCCATCGACGCCGTCCTGGAGACCGGCACCAGTGTTGCCGAGGTCGCCA 130795

Qy       2225 TGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACC 2284
Db     130796 TGGCGTACTCCGGTGACCTGTCCAATCCGGGGGAGAAGCTCTACACCCTGGACTACTACC 130855

Qy       2285 TAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGG 2344
Db     130856 TGAACCTGGCCGAGCAGATCGTCGACTCCGGTGCACACATCCTGGCCATCAAGGACATGG 130915

Qy       2345 CTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCG 2404
Db     130916 CCGGCCTGCTGCGCCCGGCCGCGGCGACCAAACTGGTCACCGCCCTGCGCCGTGAATTCG 130975

Qy       2405 ATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTG 2464
Db     130976 ACCTGCCCGTGCATGTCCACACCCACGACACCGCCGGCGGTCAGCTGGCCACCTACCTGG 131035

Qy       2465 CTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCA 2524
Db     131036 CCGCCGCCAACGCCGGGGCCGATGCCGTCGACGCCGCCTCCGCACCCCTGTCCGGTACCA 131095

Qy       2525 CCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCG 2584
Db     131096 CCTCCCAGCCGTCGATGTCCGCTCTGGTTGCCGCGTTTGCGCACACCCGACGCGACACCG 131155

Qy       2585 GTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGT 2644
Db     131156 GCCTCAACCTGCAGGCCGTCTCCGACCTGGAACCGTACTGGGAGGCGGTCCGCGGACTGT 131215

Qy       2645 ACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCC 2704
Db     131216 ACCTGCCGTTTGAATCCGGCACCCCGGGCCCGACCGGACGCGTTTACCGCCACGAGATCC 131275

Qy       2705 CAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTT 2764
Db     131276 CCGGCGGTCAGCTGTCCAACCTGCGTGCCCAGGCCGTTGCACTGGGTCTGGCCGACCGCT 131335

Qy       2765 TCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGG 2824
Db     131336 TCGAGCTCATCGAGGACTACTACGCGGCCGTCAACGAGATGCTGGGTCGTCCGACCAAGG 131395

Qy       2825 TCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGG 2884
Db     131396 TCACCCCGTCCTCCAAGGTTGTCGGTGACCTCGCACTGCACCTCGTCGGTGCCGGTGTGA 131455

Qy       2885 ATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGT 2944
Db     131456 GCCCGGAGGATTTCGCCGCCGATCCGCAGAAGTACGACATCCCCGATTCGGTCATCGCCT 131515

Qy       2945 TCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCG 3004
Db     131516 TCCTCCGCGGCGAACTGGGTACCCCTCCCGGTGGCTGGCCCGAACCGCTGCGCACCCGTG 131575

Qy       3005 CACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGG 3064
Db     131576 CACTCGAGGGTCGCTCCCAGGGTAAGGCCCCGCTGGCGGAGATCCCCGCCGAGGAGCAGG 131635

Qy       3065 CGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCC 3124
Db     131636 CCCACCTGGATTCCGATGATTCCGCGGAGCGTCGCGGCACCCTCAACCGCCTGCTGTTCC 131695

Qy       3125 CGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCCTCTGCGCTGG 3184
Db     131696 CGAAGCCGACCGAGGAGTTCCTTGAGCACCGTCGCCGCTTCGGCAACACCTCCGCCCTGG 131755

Qy       3185 ATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAG 3244
```

```
Db     131756 ATGACCGCGAGTTCTTCTACGGCTTGAAGGAGGGACGTGAGGAGCTGATCCGACTGACCG 131815
Qy       3245 ATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTA 3304
Db     131816 GTGTGTCCACCCCGATGGTGGTCCGCCTGGATGCGGTGTCCGAACCGGATGACAAAGGCA 131875
Qy       3305 TGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCT 3364
Db     131876 TGCGCAACGTGGTGGTCAACGTCAACGGCCAGATCCGCCCGATCAAGGTGCGCGACCGTT 131935
Qy       3365 CCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTG 3424
Db     131936 CCGTGGAGTCCGTCACCGCCACCGCGGAGAAGGCCGATGCCACCAACAAGGGCCATGTCG 131995
Qy       3425 CTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTG 3484
Db     131996 CCGCACCATTCGCCGGTGTGGTCACCGTGACCGTCGCCGAGGGTGATGAGATCAAGGCTG 132055
Qy       3485 GAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTG 3544
Db     132056 GCGACGCCGTGGCCATCATTGAGGCCATGAAGATGGAGGCCACCATCACCGCGCCTGTCG 132115
Qy       3545 ACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGA 3604
Db     132116 ACGGTGTCATCGACCGCGTCGTGGTGCCCGCCGCCACCAAGGTCGAGGGCGGCGACCTCA 132175
Qy       3605 TCGTCGTCGTTTCCTA 3620
Db     132176 TCGTGGTCGTGTCCTA 132191
```

RESULT 15
BD094153
LOCUS BD094153 4013 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes for heat resistant enzymes of amino acid biosynthetic pathway derived from thermophilic coryneform bacteria.
ACCESSION BD094153
VERSION BD094153.1 GI:22639741
KEYWORDS WO 0125447-A/11.
SOURCE Corynebacterium thermoaminogenes
ORGANISM Corynebacterium thermoaminogenes
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 4013)
AUTHORS Hirano,S., Nonaka,G., Matsuzaki,Y., Akiyoshi,N., Nakamura,K., Kimura,E., Osumi,T., Matsui,K., Kawahara,Y., Kurahashi,O., Nakamatsu,T. and Sugimoto,S.
TITLE Genes for heat resistant enzymes of amino acid biosynthetic pathway derived from thermophilic coryneform bacteria
JOURNAL Patent: WO 0125447-A 11 12-APR-2001;
AJINOMOTO CO INC,SEIKO HIRANO,GEN NONAKA,YUMI MATSUZAKI, NAOKI AKIYOSHI, KANAE NAKAMURA,EIICHIRO KIMURA,TSUYOSHI OSUMI,KAZUHIKO MATSUI, YOSHIO KAWAHARA,OSAMU KURAHASHI,TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO
COMMENT OS Corynebacterium thermoaminogenes
PN WO 0125447-A/11
PD 12-APR-2001
PF 04-OCT-2000 WO 2000JP006913
PR 04-OCT-1999 JP 99P 282716,01-NOV-1999 JP 99P 311147 PR 21-APR-2000 JP 00P 120687
PI SEIKO HIRANO,GEN NONAKA,YUMI MATSUZAKI,NAOKI AKIYOSHI,KANAE
PI NAKAMURA,
PI EIICHIRO KIMURA,TSUYOSHI OSUMI,KAZUHIKO MATSUI,YOSHIO KAWAHARA,
PI OSAMU KURAHASHI,TSUYOSHI NAKAMATSU,SHINICHI SUGIMOTO PC C12N15/60,C12N15/54,C12N15/53,C12N15/31,C12N15/56,C12N9/88, PC C12N9/12,
PC C12N9/04,C07K14/34,C12N9/26,C12N13/04
CC
FH Key Location/Qualifiers
FT CDS (319). .(3735).
FEATURES Location/Qualifiers
source 1. .4013

/organism="Corynebacterium thermoaminogenes"
/mol_type="genomic DNA"
/db_xref="taxon:161484"
ORIGIN

Query Match 63.0%; Score 2279.6; DB 6; Length 4013;
Best Local Similarity 77.7%; Pred. No. 0;
Matches 2808; Conservative 0; Mismatches 799; Indels 9; Gaps 4;

```
Qy          7 GGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGGTCAGGG 66
Db        129 GGGGTCTCGGGCGGGGGGCATTCTTTTCACGGCAAGGTGGTGAAATTCCGCAGGTCACTC 188
Qy         67 GAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGGAATCGG 126
Db        189 CCCGGCCGGCGGTAGAGAACGGAGCGAAAACGGAAAGCAATACGTGGTTTTCCGGACTGG 248
Qy        127 GGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTT--GGCGGTCTCTTGTTGAAAG 184
Db        249 CCGTTACGATGTTCTGAAG-AGTGACTGCCATCACCCAACAGGCTGGTCCTCGTCGAAAG 307
Qy        185 GAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCT 244
Db        308 GAACAAAAACTGTGGTTACAACA---ACACCCTCCACGCTGCCGGCGTTCAAAAAGATCC 364
Qy        245 TGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTG 304
Db        365 TGGTGGCCAACCGAGGTGAAATCGCGGTGCGAGCATTCCGCGCCGCCTACGAGACCGGGG 424
Qy        305 CAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTT 364
Db        425 CCGCAACCGTGGCCATCTACCCCCGGGAGGACCGTGGCTCCTTCCACCGCTCCTTCGCCT 484
Qy        365 CTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATG 424
Db        485 CCGAGGCGGTGAGGATCGGAACCGAGGGCTCACCCGTCAAGGCGTACCTCGATATTGATG 544
Qy        425 AAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCC 484
Db        545 AGATCATCAACGCCGCCAAGAAGGTGAAAGCGGACGCGGTCTACCCGGGGTATGGTTTCC 604
Qy        485 TGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCC 544
Db        605 TTTCGGAAAATGCCCAGCTCGCGCGTGAATGCGCGGAGAACGGCATTACCTTCATCGGTC 664
Qy        545 CAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGA 604
Db        665 CCACCCCGGAGGTGCTCGACCTCACGGGCGACAAGTCCAAGGCTGTGTCCGCCGCGAAGA 724
Qy        605 AGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTA 664
Db        725 AGGCCGGGCTGCCGGTGCTGGCGGAATCCACCCCCAGCACCGACATCGATGAGATCGTCA 784
Qy        665 AAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGAC 724
Db        785 AGAGTGCCGAGGGGCAGACCTACCCGATCTTCGTCAAGGCCGTCGCAGGTGGTGGCGGGC 844
Qy        725 GCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTC 784
Db        845 GTGGTATGCGGTTCGTCGAGAAGCCCGAGGACCTGCGTGAGCTGGCCAGGGAGGCCTCCC 904
Qy        785 GTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACC 844
Db        905 GCGAGGCGGAGGCCGCTTTCGGTGACGGATCCGTCTACGTCGAACGGGCCGTGATCAAAC 964
Qy        845 CTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATG 904
Db        965 CCCAGCACATCGAGGTGCAGATCCTCGGTGATCACACCGGCGATGTCATCCACCTGTATG 1024
Qy        905 AACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGC 964
Db       1025 AACGCGACTGTTCCCTGCAGCGCCGCCACCAGAAGGTCGTGGAGATCGCACCTGCCCAGC 1084
```

```
Qy    965 ATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCA 1024
Db   1085 ACCTCGACCCGGAGCTGCGCGACCGCATCTGTGCCGATGCCGTGAAGTTCTGCAAATCCA 1144

Qy   1025 TTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCT 1084
Db   1145 TCGGATACCAGGGCGCCGGCACCGTGGAGTTCCTCGTCGACGAGGCGGGCAACCACGTCT 1204

Qy   1085 TCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGG 1144
Db   1205 TCATTGAGATGAACCCCCGCATCCAGGTGGAACACACCGTGACCGAGGAGGTCACCTCCG 1264

Qy   1145 TGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTC 1204
Db   1265 TCGACCTGGTCAAGGCGCAGATGCACCTGGCCGCCGGTGCCACCCTGAAGGAACTGGGCC 1324

Qy   1205 TGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAG 1264
Db   1325 TGACCCAGGACAAGATCACCACCCACGGTGCCGCCCTGCAGTGCCGCATCACCACGGAGG 1384

Qy   1265 ATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCG 1324
Db   1385 ACCCGTCCAACAACTTCCGGCCCGACACCGGTGTGATCACCGCCTACCGCTCCCCGGGTG 1444

Qy   1325 GAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTG 1384
Db   1445 GTGCGGGTGTGCGTCTCGACGGCGCAGCCCAGCTCGGCGGCGAGATCACCGCACATTTCG 1504

Qy   1385 ACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTG 1444
Db   1505 ATTCCATGCTGGTCAAGATGACCTGCCGCGGTTCCGATTTCGAGACCGCCGTGTCCCGAG 1564

Qy   1445 CACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGC 1504
Db   1565 CCCAGCGCGCCCTGGCGGAGTTCAACGTCTCCGGCGTGGCCACCAACATCGGCTTCCTGC 1624

Qy   1505 GTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCG 1564
Db   1625 GTGCGCTGCTGCGCGAGGAAGACTTCACCAAGAGGCGCATCGACACCGGCTTCATCGGCT 1684

Qy   1565 ATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATT 1624
Db   1685 CCCACCAGCACCTGCTCCAGGCCCCACCGGCCGACGATGAGCAGGGGCGGATCCTGGAAT 1744

Qy   1625 ACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTC 1684
Db   1745 ACCTGGCGGATGTCACCGTGAACAAACCCCACGGTGAACGCCC---CGAGACAGCCCGTC 1801

Qy   1685 CTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCC 1744
Db   1802 CGATAGAGAAGCTGCCCGAGGTGGAGAACATCCCGCTGCCACGCGGCTCCCGCGACCGCC 1861

Qy   1745 TGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAG 1804
Db   1862 TGAAGCAGCTCGGCCCGGAGGGTTTCGCCCGCGATCTGCGCGAACAGGATGCCCTGGCCG 1921

Qy   1805 TTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCAT 1864
Db   1922 TCACCGACACCACCTTCCGCGATGCCCACCAGTCCCTCCTGGCCACCCGCGTGCGCTCCT 1981

Qy   1865 TCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGG 1924
Db   1982 TCGCGCTGACCCCGGCGGCGCGCGCCGTCGCAAAGCTCACCCCCGAGCTGCTGTCGGTGG 2041

Qy   1925 AGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGG 1984
Db   2042 AGGCCTGGGGCGGTGCCACCTACGACGTGGCCATGCGCTTCCTCTTCGAGGATCCGTGGG 2101

Qy   1985 ACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCG 2044
Db   2102 CACGCCTGGATGAGCTGCGTGAGGCGATGCCGAATGTGAACATCCAGATGCTGCTGCGTG 2161

Qy   2045 GCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGG 2104

Db   2162 GTCGCAACACCGTCGGGTACACCCCGTACCCCGATTCGGTGTGCCGCGCGTTTGTGCAGG 2221

Qy   2105 AAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCC 2164
Db   2222 AGGCCGCCAAGTCCGGTGTGGACATCTTCCGCATCTTCGACGCGCTCAACGACATCTCCC 2281

Qy   2165 AGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTA 2224
Db   2282 AGATGCGCCCGGCCATCGACGCCGTCCTGGAGACCGGCACCAGTGTTGCCGAGGTCGCCA 2341

Qy   2225 TGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACC 2284
Db   2342 TGGCGTACTCCGGTGACCTGTCCAATCCGGGGGAGAAGCTCTACACCCTGGACTACTACC 2401

Qy   2285 TAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGG 2344
Db   2402 TGAACCTGGCCGAGCAGATCGTCGACTCCGGTGCACACATCCTGGCCATCAAGGACATGG 2461

Qy   2345 CTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCG 2404
Db   2462 CCGGCCTGCTGCGCCGCGCCGCGGCGCCCAAACTGGTCACCGCCCTGCGCCGTGAATTCG 2521

Qy   2405 ATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTG 2464
Db   2522 ACCTGCCCGTGCATGTCCACACCCACGACACCGCCGGCGGTCAGCTGGCCACCTACCTGG 2581

Qy   2465 CTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCA 2524
Db   2582 CCGCCGCCAACGCCGGGGCCGATGCCGTCGACGCCGCCTCCGCACCCCTGTCCGGTACCA 2641

Qy   2525 CCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCG 2584
Db   2642 CCTCCCAGCCGTCGATGTCCGCTCTGGTTGCCGCGTTTGCGCACACCCGACGCGACACCG 2701

Qy   2585 GTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGT 2644
Db   2702 GCCTCAACCTGCAGGCCGTCTCCGACCTGGAACCGTACTGGGAGGCGGTCCGCGGACTGT 2761

Qy   2645 ACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCC 2704
Db   2762 ACCTGCCGTTTGAATCCGGCACCCCGGGCCCGACCGGACGCGTTTACCGCCACGAGATCC 2821

Qy   2705 CAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTT 2764
Db   2822 CCGGCGGTCAGCTGTCCAACCTGCGTGCCCAGGCCGTTGCACTGGGTCTGGCCGACCGCT 2881

Qy   2765 TCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGG 2824
Db   2882 TCGAGCTCATCGAGGACTACTACGCGGCCGTCAACGAGATGCTGGGTCGTCCGACCAAGG 2941

Qy   2825 TCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGG 2884
Db   2942 TCACCCCGTCCTCCAAGGTTGTCGGTGACCTCGCACTGCACCTCGTCGGTGCCGGTGTGA 3001

Qy   2885 ATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGT 2944
Db   3002 GCCCGGAGGATTTCGCCGCCGATCCGCAGAAGTACGACATCCCCGATTCGGTCATCGCCT 3061

Qy   2945 TCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCG 3004
Db   3062 TCCTCCGCGGCGAACTGGGTACCCCTCCCGGTGGCTGGCCCGAACCGCTGCGCACCCGTG 3121

Qy   3005 CACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGG 3064
Db   3122 CACTCGAGGGTCGCTCCCAGGGTAAGGCCCCGCTGGCGGAGATCCCCGCCGAGGAGCAGG 3181

Qy   3065 CGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCC 3124
Db   3182 CCCACCTGGATTCCGATGATTCCGCGGAGCGTCGCGGCACCCTCAACCGCCTGCTGTTCC 3241

Qy   3125 CGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGG 3184
```

```
Db       3242 CGAAGCCGACCGAGGAGTTCCTTGAGCACCGTCGCCGCTTCGGCAACACCTCCGCCCTGG 3301

Qy       3185 ATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAG 3244
              |||| || || |||||||||||| ||   || || || |||    |||||||| ||| | |
Db       3302 ATGACCGCGAGTTCTTCTACGGCTTGAAGGAGGGACGTGAGGAGCTGATCCGACTGACCG 3361

Qy       3245 ATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTA 3304
              ||||  ||||||  || | || ||||||||||||| | || || || || || || || | |
Db       3362 GTGTGTCCACCCCGATGGTGGTCCGCCTGGATGCGGTGTCCGAACCGGATGACAAAGGCA 3421

Qy       3305 TGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCT 3364
              ||||||| || |||| |||||||||||||||||||||| ||    ||||| ||||| |
Db       3422 TGCGCAACGTGGTGGTCAACGTCAACGGCCAGATCCGCCCGATCAAGGTGCGCGACCGTT 3481

Qy       3365 CCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTG 3424
              |||| ||||| |||||||| ||||| || ||||| ||| || |||||||||||||||| |
Db       3482 CCGTGGAGTCCGTCACCGCCACCGCGGAGAAGGCCGATGCCACCAACAAGGGCCATGTCG 3541

Qy       3425 CTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTG 3484
              | |||||||||||| ||||| |||||||||| || || || ||||||||| |||||||||
Db       3542 CCGCACCATTCGCCGGTGTGGTCACCGTGACCGTCGCCGAGGGTGATGAGATCAAGGCTG 3601

Qy       3485 GAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTG 3544
              | || || || || ||||| ||||| |||||||||| || || ||||| ||  |||| |
Db       3602 GCGACGCCGTGGCCATCATTGAGGCCATGAAGATGGAGGCCACCATCACCGCGCCTGTCG 3661

Qy       3545 ACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGA 3604
              ||||    ||||| ||||| ||||| || || || || ||||| || || |||||| | |
Db       3662 ACGGTGTCATCGACCGCGTCGTGGTGCCCGCCGCCACCAAGGTCGAGGGCGGCGACCTCA 3721

Qy       3605 TCGTCGTCGTTTCCTA 3620
              |||| ||||| |||||
Db       3722 TCGTGGTCGTGTCCTA 3737
```

Search completed: March 24, 2004, 00:13:39
Job time : 13746 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 17:20:33 ; Search time 1295 Seconds
(without alignments)
11878.552 Million cell updates/sec

Title: US-10-045-072-1
Perfect score: 3621
Sequence: 1 tggggcggggttagatcctg.....tgatcgtcgtcgtttcctaa 3621

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3621	100.0	3621	3	AAA47533	Aaa47533 Sequence
	2	3621	100.0	3621	5	AAF32165	Aaf32165 Corynebac
	3	3621	100.0	3621	6	AAD42046	Aad42046 Corynebac
	4	3621	100.0	3621	8	ACA62133	Aca62133 DNA encod
	5	3621	100.0	349980	5	AAH68526	Aah68526 C glutami
	6	3576.4	98.8	3728	2	AAX24102	Aax24102 C. glutam
	7	3420	94.4	3420	5	AAH65730	Aah65730 C glutami
	8	3398.8	93.9	3474	6	ABK52832	Abk52832 Corynebac
	9	2279.6	63.0	4013	5	AAF87437	Aaf87437 Corynebac
	10	1927.4	53.2	3423	7	ACA29831	Aca29831 Prokaryot
	11	1708	47.2	1719	4	AAF71419	Aaf71419 Corynebac
	12	1395	38.5	1406	4	AAF71420	Aaf71420 Corynebac
	13	1361.6	37.6	3381	7	ACA37896	Aca37896 Prokaryot
	14	1294.4	35.7	3384	7	ACA40756	Aca40756 Prokaryot
c	15	1294.4	35.7	110000	4	AAI99682_33	Continuation (34 o
c	16	1294.4	35.7	110000	4	AAI99683_33	Continuation (34 o
	17	1292.2	35.7	3381	7	ACA38630	Aca38630 Prokaryot
	18	1249.6	34.5	32155	9	ADB74252	Adb74252 Mycobacte
	19	1033	28.5	1083	4	AAF71418	Aaf71418 Corynebac
	20	1033	28.5	1083	4	AAF71417	Aaf71417 Corynebac
	21	907	25.0	939	4	AAF71415	Aaf71415 Corynebac
	22	907	25.0	939	4	AAF71416	Aaf71416 Corynebac
	23	681	18.8	3945	6	ABK63659	Abk63659 Rat seque
	24	681	18.8	3945	7	ABT41829	Abt41829 Toxicity
	25	681	18.8	3945	9	ADB59092	Adb59092 Toxicity-
	26	681	18.8	3945	9	ADB52590	Adb52590 Primary r
	27	677.4	18.7	4017	6	ABN96532	Abn96532 Gene #303
	28	668.4	18.5	4138	4	ABL19577	Abl19577 Drosophil
	29	668.4	18.5	4308	4	ABL19575	Abl19575 Drosophil
	30	668.4	18.5	4369	4	ABL02789	Abl02789 Drosophil
	31	623.4	17.2	4034	4	ABL20985	Abl20985 Drosophil
	32	598	16.5	4152	6	ABQ70879	Abq70879 Listeria
	33	586	16.2	3468	8	ADB08195	Adb08195 Alloiococ
	34	586	16.2	110000	8	ADB12064_05	Continuation (6 of
	35	582.4	16.1	110000	6	ABA90521_06	Continuation (7 of
	36	577.2	15.9	3441	7	ACA36434	Aca36434 Prokaryot
	37	577.2	15.9	110000	6	ABA03041_10	Continuation (11 o
	38	573.6	15.8	3462	4	AAS01509	Aas01509 Bacillus
	39	565	15.6	110000	6	ABQ69245_10	Continuation (11 o
	40	565	15.6	110000	6	ABQ67197_09	Continuation (10 o
	41	522	14.4	3444	7	ACA22812	Aca22812 Prokaryot
	42	514	14.2	3429	4	AAS53072	Aas53072 Enterococ
	43	509.2	14.1	3426	7	ACA18474	Aca18474 Prokaryot
c	44	500.4	13.8	20072	2	AAX13026	Aax13026 Enterococ
c	45	500.4	13.8	20072	6	ABS98821	Abs98821 Enterococ

ALIGNMENTS

RESULT 1
AAA47533
ID AAA47533 standard; DNA; 3621 BP.
XX
AC AAA47533;
XX
DT 20-OCT-2000 (first entry)
XX
DE Sequence encoding pyruvate carboxylase of C. glutamicum.
XX
KW Pyruvate carboxylase; expression; amino acid biosynthesis; lysine;
KW glutamic acid; oxaloacetate; fermentation; biosynthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS 199
FT /*tag= a
FT /product= "Pyruvate carboxylase"
FT /transl_except= (pos:199. .202, aa:Met)
XX
PN WO200039305-A1.
XX
PD 06-JUL-2000.
XX
PF 23-DEC-1998; 98WO-US027301.
XX
PR 23-DEC-1998; 98WO-US027301.
XX
PA (SINS/) SINSKEY A J.
PA (LESS/) LESSARD P A.
PA (WILL/) WILLIS L B.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI; 2000-465746/40.
DR P-PSDB; AAB01436.
XX
PT Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT carboxylase useful for industrial fermentation processes comprises a
PT specific nucleotide sequence.
XX
PS Claim 3; Fig 1; 51pp; English.
XX
CC The pyruvate carboxylase of Corynebacterium glutamicum can be used for
CC producing amino acids, preferably lysine and glutamic acid in industrial

CC fermentations and for replenishing oxaloacetate consumed for biosynthesis
CC during growth. By incorporating the pyruvate carboxylase gene in
CC expression vectors levels of expression can be 2 - 20 fold higher than in
CC Corynebacterium glutamicum
XX
SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 3; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920

Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980

Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
```

```
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100

Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160

Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220

Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280

Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340

Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400

Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460

Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520

Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580

Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640

Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700

Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760

Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820

Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880

Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940

Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060

Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120

Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180

Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240

Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300

Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360

Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420

Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480

Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540

Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600

Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 2
AAF32165
ID AAF32165 standard; DNA; 3621 BP.
XX
AC AAF32165;
XX
DT 12-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase coding sequence.
XX
KW Pyruvate carboxylase; anaplerotic pathway; industrial fermentation;
KW oxaloacetate; ds.
XX
OS Corynebacterium glutamicum.
XX
PN US6171833-B1.
XX
PD 09-JAN-2001.
XX
PF 23-DEC-1998; 98US-00220081.
XX
PR 23-DEC-1998; 98US-00220081.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI; 2001-122330/13.
DR P-PSDB; AAB67129.
XX
PT Novel nucleic acid encoding pyruvate carboxylase from Corynebacterium
PT glutamicum, for replenishing oxaloacetate consumed during lysine and
PT glutamic acid production in industrial fermentations.
XX
PS Claim 2; Col 23-30; 29pp; English.
XX

CC The present invention provides the protein and coding sequences of the
CC Corynebacterium glutamicum pyruvate carboxylase protein. This is an
CC enzyme in the anaplerotic pathway. It can be used in the replenishment of
CC oxaloacetate consumed during lysine and glutamic acid production in
CC industrial fermentation
XX
SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 5; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920

Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980

Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 3
AAD42046
ID AAD42046 standard; DNA; 3621 BP.
XX
AC AAD42046;
XX
DT 04-NOV-2002 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase DNA.
XX
KW Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;
KW oxaloacetate; growth; enzyme; gene; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key   Location/Qualifiers
FT CDS   199. .3621
FT   /*tag= a
FT   /product= "Pyruvate carboxylase protein"
FT   /transl_except= (pos:199. .201, aa:Met)
XX
PN US6403351-B1.
XX
PD 11-JUN-2002.
XX
PF 03-OCT-2000; 2000US-00677575.
XX
PR 23-DEC-1998; 98US-00220081.
XX
PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI; 2002-536037/57.
DR P-PSDB; AAE25601.

XX
PT Novel pyruvate carboxylase polypeptide, useful for replenishing
PT oxaloacetate consumed for biosynthesis during growth, or lysine and
PT glutamic acid production in industrial fermentation.
XX
PS Disclosure; Col 23-30; 28pp; English.
XX
CC The present invention relates to novel pyruvate carboxylase proteins and
CC polynucleotides encoding such proteins. Sequences of the invention are
CC important anaplerotic enzymes for replenishing oxaloacetate consumed for
CC biosynthesis during growth, or lysine and glutamic acid production in
CC industrial fermentation. The present sequence is C. glutamicum pyruvate
CC carboxylase DNA
XX
SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 6; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800

Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860

Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
```

```
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 4
ACA62133
ID ACA62133 standard; DNA; 3621 BP.
XX
AC ACA62133;
XX
DT 25-AUG-2003 (first entry)
XX
DE DNA encoding Corynebacterium glutamicum pyruvate carboxylase.
XX
KW Pyruvate carboxylase; gene; anaplerotic enzyme; oxaloacetate;
KW biosynthesis; growth; lysine production; glutamic acid production;
KW industrial fermentation; gene; ds.
XX
OS Corynebacterium glutamicum.
XX

FH	Key	Location/Qualifiers
FT	CDS	199. .3621
FT		/*tag= a
FT		/product= "Pyruvate carboxylase"

XX
PN US2003027305-A1.
XX
PD 06-FEB-2003.
XX
PF 15-JAN-2002; 2002US-00045072.
XX

PR 23-DEC-1998; 98US-00220081.
PR 03-OCT-2000; 2000US-00677575.
XX
PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI; 2003-479542/58.
DR P-PSDB; ABU10426.
XX
PT New pyruvate carboxylase from Corynebacterium glutamicum, useful as an
PT anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
PT during growth, or for lysine or glutamic acid production in industrial
PT fermentations.
XX
PS Claim 2; Page 12-16; 29pp; English.
XX
CC The invention describes a new isolated pyruvate carboxylase polypeptide
CC having an amino acid sequence at least 95% identical to a sequence
CC comprising 1140 amino acids from Corynebacterium glutamicum, or the
CC complete amino acid sequence encoded by the cosmid clone deposited with
CC the American Type Culture Collection. The polypeptide is useful as an
CC anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
CC during growth. The polypeptide is also useful for lysine or glutamic acid
CC production in industrial fermentations. This sequence encodes
CC Corynebacterium glutamicum pyruvate carboxylase
XX
SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 8; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
```

```
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 5
AAH68526
ID AAH68526 standard; DNA; 349980 BP.
XX
AC AAH68526;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 7061.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.

XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Disclosure; SEQ ID NO 7061; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP; 79274 A; 90638 C; 98727 G; 81341 T; 0 U; 0 Other;

Query Match 100.0%; Score 3621; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105013 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 105072

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105073 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 105132

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105133 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 105192

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105193 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 105252

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105253 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 105312

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105313 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 105372

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105373 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 105432

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105433 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 105492

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105493 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 105552

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105553 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 105612

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105613 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 105672

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105673 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 105732

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105733 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 105792

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105793 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 105852

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105853 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 105912

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105913 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 105972

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105973 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 106032

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106033 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 106092

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106093 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 106152

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106153 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 106212

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106213 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 106272

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106273 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 106332

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106333 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 106392

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106393 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 106452

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106453 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 106512

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106513 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 106572
```

```
Qy      1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106573 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 106632

Qy      1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106633 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 106692

Qy      1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106693 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 106752

Qy      1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106753 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 106812

Qy      1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106813 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 106872

Qy      1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106873 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 106932

Qy      1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106933 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 106992

Qy      1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106993 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 107052

Qy      2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107053 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 107112

Qy      2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107113 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 107172

Qy      2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107173 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 107232

Qy      2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107233 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 107292

Qy      2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107293 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 107352

Qy      2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107353 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 107412

Qy      2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107413 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 107472

Qy      2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107473 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 107532

Qy      2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107533 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 107592

Qy      2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107593 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 107652

Qy      2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107653 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 107712

Qy      2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107713 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 107772

Qy      2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107773 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 107832

Qy      2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107833 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 107892

Qy      2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107893 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 107952

Qy      2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107953 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 108012

Qy      3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108013 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 108072

Qy      3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108073 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 108132

Qy      3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108133 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 108192

Qy      3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108193 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 108252

Qy      3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108253 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 108312

Qy      3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108313 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 108372

Qy      3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108373 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 108432

Qy      3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108433 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 108492

Qy      3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108493 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 108552

Qy      3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108553 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 108612

Qy      3601 TTGATCGTCGTCGTTTCCTAA 3621
             |||||||||||||||||||||
Db    108613 TTGATCGTCGTCGTTTCCTAA 108633
```

RESULT 6
AAX24102
ID AAX24102 standard; DNA; 3728 BP.
XX

```
AC   AAX24102;
XX
DT   30-JUN-1999  (first entry)
XX
DE   C. glutamicum pyruvate carboxylase genomic DNA.
XX
KW   Pyruvate carboxylase; amino acid production; lysine production;
KW   threonine production; homoserine production; glutamate production;
KW   arginine production; feed additive; condiment; pharmaceutical;
KW   fine chemical; ss.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key             Location/Qualifiers
FT   CDS             165. .3587
FT                   /*tag=  a
FT                   /product= "pyruvate carboxylase"
XX
PN   DE19831609-A1.
XX
PD   15-APR-1999.
XX
PF   14-JUL-1998;   98DE-01031609.
XX
PR   04-OCT-1997;   97DE-01043894.
XX
PA   (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI   Peters-Wendisch P,  Eikmanns B,  Sahm H;
XX
DR   WPI; 1999-245521/21.
DR   P-PSDB; AAW93971.
XX
PT   Increasing microbial production of specific amino acids by increasing
PT   activity or expression of pyruvate carboxylase.
XX
PS   Claim 16; Page 8-11; 18pp; German.
XX
CC   This invention describes the isolation of a pyruvate carboxylase from
CC   Corynebacterium glutamicum which is used in a novel method for production
CC   of lysine, threonine, homoserine, glutamate and/or arginine, variously
CC   useful as feed additives, condiments, pharmaceuticals and intermediates
CC   for fine chemicals. Increasing pyruvate carboxylase activity increases
CC   the yield of microbial production of amino acids of the asparate and/or
CC   glutamate families, e.g. about 50% more lysine, 40% more threonine and
CC   150% more homoserine are secreted into the culture medium
XX
SQ   Sequence 3728 BP; 813 A; 1059 C; 1035 G; 821 T; 0 U; 0 Other;

  Query Match             98.8%;  Score 3576.4;  DB 2;  Length 3728;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3577;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          44 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 103
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          10 GCTTGAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAAC 69

Qy         104 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          70 CGATGTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCC 129

Qy         164 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 223
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGC 189

Qy         224 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 283
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 TTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCC 249

Qy         284 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 343
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 GTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGAT 309

Qy         344 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 CATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCA 369

Qy         404 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 463
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         370 AGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCA 429

Qy         464 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 523
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         430 TTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAA 489

Qy         524 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 583
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         490 ACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTC 549

Qy         584 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 643
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         550 GCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCA 609

Qy         644 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 703
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         610 AAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGG 669

Qy         704 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 763
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         670 CAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCA 729

Qy         764 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 823
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         730 AATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATG 789

Qy         824 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 883
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         790 TCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTG 849

Qy         884 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 943
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         850 GAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTG 909

Qy         944 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 1003
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         910 TCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATG 969

Qy        1004 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1063
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         970 CAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCG 1029

Qy        1064 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1030 ATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCG 1089

Qy        1124 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1090 TGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTG 1149

Qy        1184 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1150 CAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGC 1209

Qy        1244 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1210 AGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCA 1269

Qy        1304 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1270 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTG 1329

Qy        1364 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1330 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1389

Qy        1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1483
```

```
Db      1390 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1449
Qy      1484 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1543
Db      1450 CAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCA 1509
Qy      1544 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1603
Db      1510 TCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATG 1569
Qy      1604 AGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1663
Db      1570 AGCAGGGAGGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGC 1629
Qy      1664 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1723
Db      1630 GTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGC 1689
Qy      1724 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1783
Db      1690 CACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCC 1749
Qy      1784 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1843
Db      1750 GTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGC 1809
Qy      1844 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1903
Db      1810 TTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGA 1869
Qy      1904 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1963
Db      1870 CTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTT 1929
Qy      1964 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 2023
Db      1930 TCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAA 1989
Qy      2024 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2083
Db      1990 ACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCG 2049
Qy      2084 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2143
Db      2050 TCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCG 2109
Qy      2144 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2203
Db      2110 ACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACA 2169
Qy      2204 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2263
Db      2170 CCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGC 2229
Qy      2264 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2323
Db      2230 TCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACA 2289
Qy      2324 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2383
Db      2290 TCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCA 2349
Qy      2384 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2443
Db      2350 CCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTG 2409
Qy      2444 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2503
Db      2410 GCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTT 2469
Qy      2504 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2563
Db      2470 CCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCG 2529
Qy      2564 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2623
Db      2530 CGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACT 2589
Qy      2624 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2683
Db      2590 GGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTC 2649
Qy      2684 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2743
Db      2650 GCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCG 2709
Qy      2744 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2803
Db      2710 CACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGA 2769
Qy      2804 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2863
Db      2770 TGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCC 2829
Qy      2864 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2923
Db      2830 ACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACA 2889
Qy      2924 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2983
Db      2890 TCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGC 2949
Qy      2984 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3043
Db      2950 CAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGG 3009
Qy      3044 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3103
Db      3010 AAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATA 3069
Qy      3104 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3163
Db      3070 GCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCT 3129
Qy      3164 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3223
Db      3130 TCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCG 3189
Qy      3224 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3283
Db      3190 AGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCT 3249
Qy      3284 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3343
Db      3250 CTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCC 3309
Qy      3344 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3403
Db      3310 CAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATT 3369
Qy      3404 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3463
Db      3370 CCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTG 3429
Qy      3464 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3523
Db      3430 AAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAG 3489
Qy      3524 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3583
Db      3490 CAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGA 3549
Qy      3584 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db      3550 AGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3587
```

RESULT 7
AAH65730
ID AAH65730 standard; DNA; 3420 BP.
XX
AC AAH65730;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 765.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR P-PSDB; AAG90511.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 8; SEQ ID NO 765; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 3420 BP; 732 A; 1004 C; 939 G; 745 T; 0 U; 0 Other;

Query Match 94.4%; Score 3420; DB 5; Length 3420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      199 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 258
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 60

Qy      259 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120

Qy      319 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 378
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 180

Qy      379 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 438
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 240

Qy      439 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 498
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 300

Qy      499 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 558
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 360

Qy      559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 618
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 420

Qy      619 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 678
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 480

Qy      679 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 738
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 540

Qy      739 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 798
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 600

Qy      799 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 858
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 660

Qy      859 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 918
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 720

Qy      919 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 978
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 780

Qy      979 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 1038
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 840

Qy     1039 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 1098
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 900

Qy     1099 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 1158
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 960

Qy     1159 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1218
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1020

Qy     1219 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1278
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1080

Qy     1279 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1338
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1140

Qy     1339 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1398
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1200

Qy     1399 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1458
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1260

Qy     1459 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1518
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1261 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1320
Qy       1519 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1578
Db       1321 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1380
Qy       1579 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1638
Db       1381 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1440
Qy       1639 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1698
Db       1441 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1500
Qy       1699 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1758
Db       1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1560
Qy       1759 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1818
Db       1561 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620
Qy       1819 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1878
Db       1621 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680
Qy       1879 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1938
Db       1681 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1740
Qy       1939 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1998
Db       1741 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800
Qy       1999 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 2058
Db       1801 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 1860
Qy       2059 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 2118
Db       1861 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 1920
Qy       2119 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2178
Db       1921 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980
Qy       2179 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2238
Db       1981 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2040
Qy       2239 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2298
Db       2041 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2100
Qy       2299 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2358
Db       2101 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2160
Qy       2359 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2418
Db       2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2220
Qy       2419 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2478
Db       2221 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2280
Qy       2479 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2538
Db       2281 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2340
Qy       2539 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2598
Db       2341 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2400
```

```
Qy       2599 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2658
Db       2401 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2460
Qy       2659 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2718
Db       2461 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2520
Qy       2719 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2778
Db       2521 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2580
Qy       2779 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2838
Db       2581 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2640
Qy       2839 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2898
Db       2641 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2700
Qy       2899 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2958
Db       2701 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2760
Qy       2959 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 3018
Db       2761 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 2820
Qy       3019 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 3078
Db       2821 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 2880
Qy       3079 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 3138
Db       2881 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 2940
Qy       3139 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3198
Db       2941 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3000
Qy       3199 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3258
Db       3001 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3060
Qy       3259 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3318
Db       3061 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3120
Qy       3319 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3378
Db       3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3180
Qy       3379 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3438
Db       3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3240
Qy       3439 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3498
Db       3241 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3300
Qy       3499 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3558
Db       3301 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360
Qy       3559 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3618
Db       3361 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3420
```

RESULT 8
ABK52832
ID ABK52832 standard; DNA; 3474 BP.

```
XX
AC   ABK52832;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Corynebacterium glutamicum feedback-resistant pyruvate carboxylase gene.
XX
KW   Feedback-resistant; pyruvate carboxylase enzyme; gene; ds;
KW   aspartic acid feedback inhibition resistant.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3474
FT                   /*tag=  a
FT                   /product= "Corynebacterium glutamicum feedback-resistant
FT                   pyruvate carboxylase enzyme"
FT                   /transl_except= (pos:1. .3; aa:Met)
XX
PN   WO200231158-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US031893.
XX
PR   13-OCT-2000; 2000US-0239913P.
XX
PA   (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
PI   Hanke PD;
XX
DR   WPI; 2002-463267/49.
DR   P-PSDB; AAU98050.
XX
PT   Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT   polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,
PT   L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.
XX
PS   Claim 3; Fig 1; 42pp; English.
XX
CC   The present invention relates to a new mutated, feedback-resistant
CC   pyruvate carboxylase enzyme. The invention is useful for producing an
CC   amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro), by
CC   culturing a host cell in a suitable media and separating the amino acid
CC   from the medium. The vector of the invention is useful for replacement of
CC   a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate
CC   carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic
CC   copy of the wild-type pyruvate carboxylase gene with a selectable marker
CC   gene through homologous recombination to form a first recombination
CC   strain, and replacing the selectable marker gene in the first recombinant
CC   strain, with feedback resistant pyruvate carboxylase gene through
CC   homologous recombination to form a second recombinant strain, where the
CC   homologous recombination in the above steps, occurs between the host cell
CC   and the vector. The feedback-resistant pyruvate carboxylase enzyme is
CC   resistant to feedback inhibition from aspartic acid. The present nucleic
CC   acid sequence encodes the feedback-resistant pyruvate carboxylase enzyme
CC   of the invention
XX
SQ   Sequence 3474 BP; 749 A; 1013 C; 950 G; 762 T; 0 U; 0 Other;

  Query Match             93.9%;  Score 3398.8;  DB 6;  Length 3474;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 3427;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy        148 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy        208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120

Qy        268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180

Qy        328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240

Qy        388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300

Qy        448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
              ||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||
Db        301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy        508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420

Qy        568 ACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 480

Qy        628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540

Qy        688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
              |||||||||| |||||||||||||||||||||||||||||||||| |||||| |||||
Db        541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCA 600

Qy        748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
              |||||||||| |||||||| ||||||||||||||||||||||||||||| ||||||
Db        601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660

Qy        808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
              || ||  ||||||||||||||||||||||||||||||| ||||| ||||||||||||||
Db        661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720

Qy        868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780

Qy        928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840

Qy        988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900

Qy       1048 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC 960

Qy       1108 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020

Qy       1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080

Qy       1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140

Qy       1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200

Qy       1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260

Qy      1408 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320

Qy      1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1380

Qy      1528 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1587
             |||||||||||||||||||||||||||||  ||  |  |||||||||  ||||||||||||||||
Db      1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT 1440

Qy      1588 CCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647
             ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500

Qy      1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
             |||||||||||||||||||||||||||||||||||||||  ||  ||||||||||||||  ||||||
Db      1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560

Qy      1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1620

Qy      1768 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680

Qy      1828 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740

Qy      1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1800

Qy      1948 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 2007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860

Qy      2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920

Qy      2068 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980

Qy      2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2187
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040

Qy      2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100

Qy      2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160

Qy      2308 AAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2367
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db      2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220

Qy      2368 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280

Qy      2428 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487
             ||||||||||||||||||||  ||||  ||||||||||||||||||||||||||||||||||||
Db      2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340

Qy      2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2400

Qy      2548 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2401 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460

Qy      2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2667
             |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db      2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520

Qy      2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580

Qy      2728 CGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC 2787
             |||||||||||||||||||||||||||||  |||||  |||||  ||||||||||||||||||
Db      2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640

Qy      2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700

Qy      2848 GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2907
             ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||  ||
Db      2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760

Qy      2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2820

Qy      2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 3027
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db      2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 2880

Qy      3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 3087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 2940

Qy      3088 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3147
             ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000

Qy      3148 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC 3207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060

Qy      3208 CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3267
             ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120

Qy      3268 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180

Qy      3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240

Qy      3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300

Qy      3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3507
             ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360

Qy      3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 3567
             ||||||||||||||||||||||||||||||||||||||||||||||||||  ||  ||||||||||
Db      3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420
```

```
Qy       3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3474


RESULT 9
AAF87437
ID   AAF87437 standard; DNA; 4013 BP.
XX
AC   AAF87437;
XX
DT   09-JUL-2001  (first entry)
XX
DE   Corynebacterium thermoaminogenes pc nucleotide sequence.
XX
KW   Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW   thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD; pdhA; pc;
KW   ppc; acn; icd; lpd; odhA; ds.
XX
OS   Corynebacterium thermoaminogenes.
XX
FH   Key             Location/Qualifiers
FT   CDS             319. .3738
FT                   /*tag=  a
FT                   /product= "pc protein"
XX
PN   WO200125447-A1.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-JP006913.
XX
PR   04-OCT-1999;   99JP-00282716.
PR   01-NOV-1999;   99JP-00311147.
PR   21-APR-2000; 2000JP-00120687.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Hirano S,  Nonaka G,  Matsuzaki Y,  Akiyoshi N,  Nakamura K;
PI   Kimura E,  Osumi T,  Matsui K,  Kawahara Y,  Kurahashi O,  Nakamatsu T;
PI   Sugimoto S;
XX
DR   WPI; 2001-300170/31.
DR   P-PSDB; AAB83180.
XX
PT   Proteins and their DNA useful for microbial production of L-amino acids.
XX
PS   Claim 34; Page 126-132; 215pp; Japanese.
XX
CC   The present sequence is provided in a specification relating to genes
CC   encoding thermophilic amino acid biosynthesis system enzymes of the
CC   thermotolerant bacterium Corynebacterium thermoaminogenes. The novel
CC   proteins retain at least 30% isocitrate ligase activity after heating at
CC   50oC for 5 minutes. DNA fragments encoding the enzymes were isolated from
CC   a Corynebacterium thermoaminogenes chromosomal DNA plasmid library by
CC   PCR. The DNA may be used for developing strains of amino acid producing
CC   microorganisms
XX
SQ   Sequence 4013 BP; 735 A; 1376 C; 1234 G; 668 T; 0 U; 0 Other;

  Query Match             63.0%;  Score 2279.6;  DB 5;  Length 4013;
  Best Local Similarity   77.7%;  Pred. No. 0;
  Matches 2808;  Conservative    0;  Mismatches  799;  Indels    9;  Gaps    4;

Qy          7 GGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGGTCAGGG 66
              |||||     | |||||  || |||      |   |  |||| |  ||||||||
Db        129 GGGGTCTCGGGCGGGGGGCATTCTTTTCACGGCAAGGTGGTGAAATTCCGCAGGTCACTC 188

Qy         67 GAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGGAATCGG 126
                | | | | | | |    ||||||  ||| | ||   || ||    | |  ||
Db        189 CCCGGCCGGCGGTAGAGAACGGAGCGAAAACGGAAAGCAATACGTGGTTTTCCGGACTGG 248

Qy        127 GGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTT--GGCGGTCTCTTGTTGAAAG 184
-

Db        249 CCGTTACGATGTTCTGAAG-AGTGACTGCCATCACCCAACAGGCTGGTCCTCGTCGAAAG 307

Qy        185 GAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCT 244
              |||  ||  |||  |  ||  |  ||   |||  |  ||  |||||  ||  ||  ||||||||||||
Db        308 GAACAAAAACTGTGGTTACAACA---ACACCCTCCACGCTGCCGGCGTTCAAAAAGATCC 364

Qy        245 TGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTG 304
              ||||  ||  |||||  ||  ||||||||||||  ||  ||  |||||  ||  ||    |||  |||||  |
Db        365 TGGTGGCCAACCGAGGTGAAATCGCGGTGCGAGCATTCCGCGCCGCCTACGAGACCGGGG 424

Qy        305 CAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTT 364
              |  ||  ||  ||  ||  ||||||||  ||  ||  ||  ||  ||||||||||  ||  ||  |
Db        425 CCGCAACCGTGGCCATCTACCCCCGGGAGGACCGTGGCTCCTTCCACCGCTCCTTCGCCT 484

Qy        365 CTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATG 424
              |  ||  ||  ||  |  ||  ||  |||||  ||||||||  ||||||||||||||  ||  ||  ||||
Db        485 CCGAGGCGGTGAGGATCGGAACCGAGGGCTCACCCGTCAAGGCGTACCTCGATATTGATG 544

Qy        425 AAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCC 484
              |  ||  |||    ||  ||  ||  ||  ||  |||||  ||  ||  |  |||||||||  ||  ||  ||||
Db        545 AGATCATCAACGCCGCCAAGAAGGTGAAAGCGGACGCGGTCTACCCGGGGTATGGTTTCC 604

Qy        485 TGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCC 544
              |  ||  ||||||||||||||  ||  ||  ||  ||  |||||  ||||||||||  ||  ||  ||  |
Db        605 TTTCGGAAAATGCCCAGCTCGCGCGTGAATGCGCGGAGAACGGCATTACCTTCATCGGTC 664

Qy        545 CAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGA 604
              |  ||||||  |||||  ||  ||  |||||  ||  ||  |||||    ||  ||  ||||||||||||
Db        665 CCACCCCGGAGGTGCTCGACCTCACGGGCGACAAGTCCAAGGCTGTGTCCGCCGCGAAGA 724

Qy        605 AGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTA 664
              ||||  ||  |||||  ||  ||||||||||||||||  ||||    ||||||||||||  |
Db        725 AGGCCGGGCTGCCGGTGCTGGCGGAATCCACCCCCAGCACCGACATCGATGAGATCGTCA 784

Qy        665 AAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGAC 724
              |  ||  ||  ||  |||||  |||||  ||  ||  ||||||||  ||  ||  |||||||||  |
Db        785 AGAGTGCCGAGGGGCAGACCTACCCGATCTTCGTCAAGGCCGTCGCAGGTGGTGGCGGGC 844

Qy        725 GCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTC 784
              |  ||||||||  ||  ||  |      ||  ||  ||  ||  |  |  ||  ||  ||||
Db        845 GTGGTATGCGGTTCGTCGAGAAGCCCGAGGACCTGCGTGAGCTGGCCAGGGAGGCCTCCC 904

Qy        785 GTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACC 844
              |  ||  ||  ||  ||  ||||||||  ||  ||  |  ||  ||  ||||||||  ||  ||||||  ||  |
Db        905 GCGAGGCGGAGGCCGCTTTCGGTGACGGATCCGTCTACGTCGAACGGGCCGTGATCAAAC 964

Qy        845 CTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATG 904
              |  |||||  ||  ||  ||||||||||  ||  ||||||||  ||  ||  ||  |  |||||  ||||
Db        965 CCCAGCACATCGAGGTGCAGATCCTCGGTGATCACACCGGCGATGTCATCCACCTGTATG 1024

Qy        905 AACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGC 964
              ||||  |||||  ||  ||||||||  ||  ||||  ||  ||  ||  ||  ||  ||  ||||
Db       1025 AACGCGACTGTTCCCTGCAGCGCCGCCACCAGAAGGTCGTGGAGATCGCACCTGCCCAGC 1084

Qy        965 ATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCA 1024
              |  |  ||  ||  ||  |||||  ||  |||||  |||||  |||||  ||  |||||||||     ||||
Db       1085 ACCTCGACCCGGAGCTGCGCGACCGCATCTGTGCCGATGCCGTGAAGTTCTGCAAATCCA 1144

Qy       1025 TTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCT 1084
              |  ||  ||||||||||||  ||  ||||||||  |||  |  |||||  ||    ||||||||||||||
Db       1145 TCGGATACCAGGGCGCCGGCACCGTGGAGTTCCTCGTCGACGAGGCGGGCAACCACGTCT 1204

Qy       1085 TCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGG 1144
              ||||  ||  |||||||||  ||  |||||||||  ||  ||||||||||||  ||  ||  ||||||    |
Db       1205 TCATTGAGATGAACCCCCGCATCCAGGTGGAACACACCGTGACCGAGGAGGTCACCTCCG 1264

Qy       1145 TGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTC 1204
              |  ||||||||  ||||||||||||  |  ||||  |||  ||||||  ||  ||||||    |
Db       1265 TCGACCTGGTCAAGGCGCAGATGCACCTGGCCGCCGGTGCCACCCTGAAGGAACTGGGCC 1324

Qy       1205 TGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAG 1264
              |||||||  ||  |||||||  ||||||||||  ||  ||||||||||||||||||||  |
```

```
Db      1325 TGACCCAGGACAAGATCACCACCCACGGTGCCGCCCTGCAGTGCCGCATCACCACGGAGG 1384
Qy      1265 ATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCG 1324
Db      1385 ACCCGTCCAACAACTTCCGGCCCGACACCGGTGTGATCACCGCCTACCGCTCCCCGGGTG 1444
Qy      1325 GAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTG 1384
Db      1445 GTGCGGGTGTGCGTCTCGACGGCGCAGCCCAGCTCGGCGGCGAGATCACCGCACATTTCG 1504
Qy      1385 ACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTG 1444
Db      1505 ATTCCATGCTGGTCAAGATGACCTGCCGCGGTTCCGATTTCGAGACCGCCGTGTCCCGAG 1564
Qy      1445 CACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGC 1504
Db      1565 CCCAGCGCGCCCTGGCGGAGTTCAACGTCTCCGGCGTGGCCACCAACATCGGCTTCCTGC 1624
Qy      1505 GTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCG 1564
Db      1625 GTGCGCTGCTGCGCGAGGAAGACTTCACCAAGAGGCGCATCGACACCGGCTTCATCGGCT 1684
Qy      1565 ATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATT 1624
Db      1685 CCCACCAGCACCTGCTCCAGGCCCCACCGGCCGACGATGAGCAGGGGCGGATCCTGGAAT 1744
Qy      1625 ACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTC 1684
Db      1745 ACCTGGCGGATGTCACCGTGAACAAACCCCACGGTGAACGCCC---CGAGACAGCCCGTC 1801
Qy      1685 CTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCC 1744
Db      1802 CGATAGAGAAGCTGCCCGAGGTGGAGAACATCCCGCTGCCACGCGGCTCCCGCGACCGCC 1861
Qy      1745 TGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAG 1804
Db      1862 TGAAGCAGCTCGGCCCGGAGGGTTTCGCCCGCGATCTGCGCGAACAGGATGCCCTGGCCG 1921
Qy      1805 TTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCAT 1864
Db      1922 TCACCGACACCACCTTCCGCGATGCCCACCAGTCCCTCCTGGCCACCCGCGTGCGCTCCT 1981
Qy      1865 TCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGG 1924
Db      1982 TCGCGCTGACCCCGGCGGCGCGCGCCGTCGCAAAGCTCACCCCCGAGCTGCTGTCGGTGG 2041
Qy      1925 AGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGG 1984
Db      2042 AGGCCTGGGGCGGTGCCACCTACGACGTGGCCATGCGCTTCCTCTTCGAGGATCCGTGGG 2101
Qy      1985 ACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCG 2044
Db      2102 CACGCCTGGATGAGCTGCGTGAGGCGATGCCGAATGTGAACATCCAGATGCTGCTGCGTG 2161
Qy      2045 GCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGG 2104
Db      2162 GTCGCAACACCGTCGGGTACACCCCGTACCCCGATTCGGTGTGCCGCGCGTTTGTGCAGG 2221
Qy      2105 AAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCC 2164
Db      2222 AGGCCGCCAAGTCCGGTGTGGACATCTTCCGCATCTTCGACGCGCTCAACGACATCTCCC 2281
Qy      2165 AGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTA 2224
Db      2282 AGATGCGCCCGGCCATCGACGCCGTCCTGGAGACCGGCACCAGTGTTGCCGAGGTCGCCA 2341
Qy      2225 TGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACC 2284
Db      2342 TGGCGTACTCCGGTGACCTGTCCAATCCGGGGGAGAAGCTCTACACCCTGGACTACTACC 2401
Qy      2285 TAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGG 2344
Db      2402 TGAACCTGGCCGAGCAGATCGTCGACTCCGGTGCACACATCCTGGCCATCAAGGACATGG 2461

Qy      2345 CTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCG 2404
Db      2462 CCGGCCTGCTGCGCCGCGCCGCGGCGCCCAAACTGGTCACCGCCCTGCGCCGTGAATTCG 2521
Qy      2405 ATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTG 2464
Db      2522 ACCTGCCCGTGCATGTCCACACCCACGACACCGCCGGCGGTCAGCTGGCCACCTACCTGG 2581
Qy      2465 CTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCA 2524
Db      2582 CCGCCGCCAACGCCGGGGCCGATGCCGTCGACGCCGCCTCCGCACCCCTGTCCGGTACCA 2641
Qy      2525 CCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCG 2584
Db      2642 CCTCCCAGCCGTCGATGTCCGCTCTGGTTGCCGCGTTTGCGCACACCCGACGCGACACCG 2701
Qy      2585 GTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGT 2644
Db      2702 GCCTCAACCTGCAGGCCGTCTCCGACCTGGAACCGTACTGGGAGGCGGTCCGCGGACTGT 2761
Qy      2645 ACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCC 2704
Db      2762 ACCTGCCGTTTGAATCCGGCACCCCGGGCCCGACCGGACGCGTTTACCGCCACGAGATCC 2821
Qy      2705 CAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTT 2764
Db      2822 CCGGCGGTCAGCTGTCCAACCTGCGTGCCCAGGCCGTTGCACTGGGTCTGGCCGACCGCT 2881
Qy      2765 TCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGG 2824
Db      2882 TCGAGCTCATCGAGGACTACTACGCGGCCGTCAACGAGATGCTGGGTCGTCCGACCAAGG 2941
Qy      2825 TCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGG 2884
Db      2942 TCACCCCGTCCTCCAAGGTTGTCGGTGACCTCGCACTGCACCTCGTCGGTGCCGGTGTGA 3001
Qy      2885 ATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGT 2944
Db      3002 GCCCGGAGGATTTCGCCGCCGATCCGCAGAAGTACGACATCCCCGATTCGGTCATCGCCT 3061
Qy      2945 TCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCG 3004
Db      3062 TCCTCCGCGGCGAACTGGGTACCCCTCCCGGTGGCTGGCCCGAACCGCTGCGCACCCGTG 3121
Qy      3005 CACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGG 3064
Db      3122 CACTCGAGGGTCGCTCCCAGGGTAAGGCCCCGCTGGCGGAGATCCCCGCCGAGGAGCAGG 3181
Qy      3065 CGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCC 3124
Db      3182 CCCACCTGGATTCCGATGATTCCGCGGAGCGTCGCGGCACCCTCAACCGCCTGCTGTTCC 3241
Qy      3125 CGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGG 3184
Db      3242 CGAAGCCGACCGAGGAGTTCCTTGAGCACCGTCGCCGCTTCGGCAACACCTCCGCCCTGG 3301
Qy      3185 ATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAG 3244
Db      3302 ATGACCGCGAGTTCTTCTACGGCTTGAAGGAGGGACGTGAGGAGCTGATCCGACTGACCG 3361
Qy      3245 ATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTA 3304
Db      3362 GTGTGTCCACCCCGATGGTGGTCCGCCTGGATGCGGTGTCCGAACCGGATGACAAAGGCA 3421
Qy      3305 TGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCT 3364
Db      3422 TGCGCAACGTGGTGGTCAACGTCAACGGCCAGATCCGCCCGATCAAGGTGCGCGACCGTT 3481
Qy      3365 CCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTG 3424
Db      3482 CCGTGGAGTCCGTCACCGCCACCGCGGAGAAGGCCGATGCCACCAACAAGGGCCATGTCG 3541
```

```
Qy       3425 CTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTG 3484
             | |||||||||||| ||||| |||||||||||| || || || ||||||||| |||||||||
Db       3542 CCGCACCATTCGCCGGTGTGGTCACCGTGACCGTCGCCGAGGGTGATGAGATCAAGGCTG 3601

Qy       3485 GAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTG 3544
             | || || || || ||||| ||||| |||||||||||| || || ||||| ||  |||| |
Db       3602 GCGACGCCGTGGCCATCATTGAGGCCATGAAGATGGAGGCCACCATCACCGCGCCTGTCG 3661

Qy       3545 ACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGA 3604
             ||||    ||||| ||||| ||||| || || || || ||||| || || |||||| | |
Db       3662 ACGGTGTCATCGACCGCGTCGTGGTGCCCGCCGCCACCAAGGTCGAGGGCGGCGACCTCA 3721

Qy       3605 TCGTCGTCGTTTCCTA 3620
             |||| ||||| |||||
Db       3722 TCGTGGTCGTGTCCTA 3737
```

RESULT 10
ACA29831
ID ACA29831 standard; DNA; 3423 BP.
XX
AC ACA29831;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #11488.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Corynebacterium diphtheriae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU25961.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 17701; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3423 BP; 773 A; 1058 C; 912 G; 680 T; 0 U; 0 Other;

  Query Match   53.2%; Score 1927.4; DB 7; Length 3423;
  Best Local Similarity 73.4%; Pred. No. 0;
  Matches 2494; Conservative   0; Mismatches 896; Indels   9; Gaps   2;

```
Qy        213 ATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGT 272
              | | || ||| |      ||||   ||||||| | || || ||||| ||||| |||||  ||
Db         27 AACCTCTACGGTAAATCCATTGTCAAAGATCCTCGTTGCTAACCGTGGCGAGATCGCCGT 86

Qy        273 CCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGA 332
              || ||||| |||||||| | |||||||||||| |||||||||||| | |||||    ||
Db         87 GCGAGCTTTTCGTGCAGCCTTTGAAACCGGTGCTGCCACGGTAGCTGTCTACCCGAACGA 146

Qy        333 AGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGG 392
              ||| ||    |||||||||||| || ||||| |||||||||||||| |||||||   | ||
Db        147 AGACCGTAATTCATTCCACCGATCGTTTGCCTCTGAAGCTGTCCTCATTGGTGAGGGCGG 206

Qy        393 CTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAA 452
              ||  | ||||| |||||||| || ||||| ||||| ||| |||| |||||| ||
Db        207 ATCTGCGGTCAAAGCGTACCTCGATATCGACGAAATCATCCGTGCTGCTAAACAAACGGG 266

Qy        453 AGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGA 512
              || || || |||||||| || |||||||||||| || ||||| |||||||||| || ||
Db        267 TGCCGACGCAATTTACCCAGGCTACGGCTTCCTTTCCGAAAACGCCCAGCTTGCTCGTGA 326

Qy        513 GTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGG 572
              ||| || |||||||||||| || ||||| || ||| | |||    || || || || |||||
Db        327 GTGCGCTGAAAACGGCATAACCTTTATCGGTCCACCTCCATCAGTGCTGGAGCTTACCGG 386

Qy        573 TGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATC 632
              |||||| |    || || |||||  ||      |||| ||| ||||      || | ||  |
Db        387 TGATAAAGCAGCTGCTGTTACCGCAGCACGCGAGGCAGGTTTGCCCACGCTGACTGAGAC 446

Qy        633 CACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCAT 692
              |  |  |  |  |  || |||  |  |  | || ||||| |||||  |
Db        447 AGAAGCAACAGACGATCCAAAGAAGCTCGCAGAGCTGACCAAGGGGCAGACCTACCCGCT 506

Qy        693 CTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGA 752
              || || || || |||||||||||||| || ||||| |||||||||||  |  || ||
Db        507 ATTCGTCAAAGCCGTTGCCGGTGGTGGTGGCCGCGGCATGCGTTTTGTCGACGCCCCCGA 566

Qy        753 TGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGG 812
              | ||    ||  ||||| ||||||| || ||||| || |  || || ||||||||||
Db        567 AAACCTAGAAAAGCTAGCAGCAGAAGCCTCCCGTGAGGCGGCTGCAGCATTCGGCGATGG 626

Qy        813 CGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGG 872
              |    || || | ||| ||||| || || || ||||||||||||| || ||||||||||
Db        627 CCGAGTCTACGCCGAGCGTGCAGTAATCAATCCTCAGCATATTGAGGTCCAGATCCTTGG 686

Qy        873 CGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCA 932
              ||  | ||||  |  |  | ||||| || || ||||| ||||| |||||||||| ||
Db        687 TGACTCCGCTGGCAACATCATCCACCTGTACGAGCGTGATTGCTCGCTGCAGCGTCGCCA 746

Qy        933 CCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCAT 992
              |||||| || || || || || |||||||||||  ||||  || |  ||| || ||    ||
```

```
Db        747 CCAAAAGGTCGTAGAGATCGCCCCAGCACAGCACTTGGCGCCCGGACTACGCGAAAAGAT 806
Qy        993 TTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGA 1052
Db        807 TTGTGCAGACGCAGTCGCGTTTGCACGCCACATCGGTTACCAAGGTGCAGGCACGGTCGA 866
Qy       1053 ATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGT 1112
Db        867 GTTCCTAGTCGATGAGGAAGGCAACCACGTCTTCATCGAGATGAACCCTCGTATCCAGGT 926
Qy       1113 TGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTT 1172
Db        927 CGAACACACCGTGACCGAAGAGGTCACCCAAGTCGACTTGGTGAAGTCCCAGATCATGAT 986
Qy       1173 GGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGG 1232
Db        987 CGCCTCCGGTGCGACACTCGAAGATCTCGGTCTGCGCCAAGAAGACATCCACACCGAAGG 1046
Qy       1233 TGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATAC 1292
Db       1047 CGCTGCCCTGCAGTGCCGTATTACTACTGAAGACCCCAACAACGGTTTCCGTCCTGACAC 1106
Qy       1293 CGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGC 1352
Db       1107 CGGAACGATTACCGCTTATCGTTCCCCAGGTGGCGCAGGTGTTCGTCTCGACGGTGCTGC 1166
Qy       1353 TCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCG 1412
Db       1167 CATGCTCGGTGGTGAGATCACCCCGAATTTCGACTCCATGCTGGTCAAGATGACCTGCCG 1226
Qy       1413 TGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGT 1472
Db       1227 TGGTGCCGATTTCGCCACCGCAGTCGCTCGCGCTCAGCGTGCACTCGCAGAATTCGTTGT 1286
Qy       1473 GTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCAC 1532
Db       1287 CTCAGGTGTTGCTACCAACATCGGATTCCTGCGTGCTTTGCTACGCGAAGAAGACTTCCA 1346
Qy       1533 TTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACC 1592
Db       1347 GCACAAGCGCATCGCCACTGGCTTCATCGCAGATCACCCATGGCTGCTGCAGGCACCACC 1406
Qy       1593 TGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCC 1652
Db       1407 TGCCGACGATGAGCCTGGCCGCATTTTGAACTACCTAGCTGATGTGACTGTGAACAAGCC 1466
Qy       1653 TCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGA 1712
Db       1467 TCATGGTCTACGCCCA---GCCGTGATCAACCCGGTAGAAAAGCTCCCAGCAGAGATCAA 1523
Qy       1713 TCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGC 1772
Db       1524 AGGTGAGCTACCACGTGGTTCCCGTGATCGCCTGCTCCAGCTCGGCCCAGAGGGATTTGC 1583
Qy       1773 TCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACA 1832
Db       1584 GCGCGCACTGCGTAAACAGGATGCGCTCGCGGTTACGGATACCACCTTCCGTGATGCACA 1643
Qy       1833 CCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGT 1892
Db       1644 TCAGTCGCTTCTCGCAACCCGCGTGCGCTCCAACACGCTGATCGACGCAGCACGGCACGT 1703
Qy       1893 CGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGT 1952
Db       1704 CGCAAAGCTCACCCCAGAGCTGCTCTCTGTAGAAGCATGGGGTGGTGCTACCTATGACGT 1763
Qy       1953 GGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGAT 2012
Db       1764 TGCTATGCGCTTCCTACATGAGGACCCCTGGGAGCGTCTCGACCACCTGCGTGAGGCCAT 1823
Qy       2013 GCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTA 2072
Db       1824 GCCCAACGTCAACATTCAGATGCTGCTGCGAGGCCGCAACACAGTCGGCTACACACCGTA 1883

Qy       2073 CCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTT 2132
Db       1884 CCCAGATTCTGTATGCAAGGCTTTTGTCGACGAAGCCGCACGCTCTGGTGTTGATATCTT 1943
Qy       2133 CCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCT 2192
Db       1944 CCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTGCT 2003
Qy       2193 GGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCC 2252
Db       2004 CGAAACCAACACCACGATTGCCGAAGTCGCAATGGCCTACTCCGGTGATCTGACCAACCC 2063
Qy       2253 AAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTC 2312
Db       2064 AAGTGAGAAGCTCTACACACTGGATTACTACCTTAAGCTTGCCGAGGAAATCGTTAAGTC 2123
Qy       2313 TGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAAC 2372
Db       2124 CGGTGCACATGTTCTAGCCATTAAAGACATGGCTGGACTGATGAAGCCAGCCGCAGCAAC 2183
Qy       2373 CAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGA 2432
Db       2184 CAAGCTGGTCACCGAACTACGACGCAACTTCGACCTGCCGGTTCATGTTCACACCCACGA 2243
Qy       2433 CACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGT 2492
Db       2244 CACTGCTGGTGGCCAGCTGGCTACGTATTGGGCTGCAGCAGCCGCTGGTGCAGATGCCGT 2303
Qy       2493 TGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGT 2552
Db       2304 CGATGGTGCCTCCGCGCCACTGTCGGGCACAACGTCGCAGCCATCGTTGTCTGCAATTGT 2363
Qy       2553 TGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCT 2612
Db       2364 GGCAGCATTCGCTAACACCTACCGCGATACGGGTCTCTCTCTGGACGCAGTCGGATCGAT 2423
Qy       2613 CGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGG 2672
Db       2424 GGAACCGTACTGGGAAGCAGTGCGCAAACTCTACGCACCCTTCGAATCCGGAACCCCAGG 2483
Qy       2673 CCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGC 2732
Db       2484 ACCAACTGGCCGCGTCTACCAGCACGAAATCCCAGGCGGCCAGCTGTCCAATCTGCGCGC 2543
Qy       2733 ACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGC 2792
Db       2544 ACAGGCAACTGCACTCGGCTTGGCAGATCGCTTCGAGCTCATCGAGGACTACTACGCAGC 2603
Qy       2793 CGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGA 2852
Db       2604 CGTTAATGAGATGCTGGGACGCCCCACCAAGGTCACCCCGTCCTCAAAGGTCGTTGGCGA 2663
Qy       2853 CCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACA 2912
Db       2664 TCTTGCACTCTACCTAGTCGGCGCAGGAGTTAACCCAGCGGACTTTGCAGCAGATCCCCA 2723
Qy       2913 AAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCC 2972
Db       2724 AAAGTACGATATTCCTGACTCCGTCATTGCCTTCCTCCGTGGTGAGCTCGGCACCCCTCC 2783
Qy       2973 AGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGC 3032
Db       2784 AGGCGGCTGGCCAGAAGAACTGCGTGCTAAGGCATTGGCAGGACGTAAGGAATCAAAGGA 2843
Qy       3033 ACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGA 3092
Db       2844 CACATTGGCCCCGCTGCCAGCAGAAGATGAAGCTCTGCTCGCAGATCGTGCGACC----- 2898
Qy       3093 ACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCA 3152
Db       2899 -GTTCGCCCAACGCTGGATCGCCTGCTCTTCCCGAAGCCTGCGGCAGAATTCGCAGAGCA 2957
```

```
Qy       3153 CCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGT 3212
             ||| || |  || ||| | ||| |    || | |||     |||||| |||| |||||
Db       2958 CCGCCGTCAGTTTGGCGATACCACCAAACTCGGTGACGCAGAATTCCTCTATGGCCTCAA 3017

Qy       3213 CGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCT 3272
              ||||||   ||  ||  | ||  ||     ||||      ||  |||| |||| || || ||
Db       3018 AGAAGGCAAAGAAACGGTCATTCGTACCGCAGACAGTAGCGTCCCAATGCTGGTGCGTCT 3077

Qy       3273 GGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGG 3332
              |||||  ||   ||||| ||||| ||||| |||||||| ||||||  ||| ||||||||
Db       3078 CGATGCCGTCGGCGAGCCCGACGAAAAGGGCATGCGCAACGTTGTGTGCAATGTCAACGG 3137

Qy       3333 CCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGA 3392
              |||||| ||||| || |  |||||||| ||| ||||| || |||||  | |  ||
Db       3138 CCAGATTCGCCCCATCCTCGTGCGTGATCGCAGCGTAGAGTCAGTGACCGCCTCGGTGGA 3197

Qy       3393 AAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGT 3452
              |||||| ||| | ||||||   || || || || |||||||||| || ||||||||  ||
Db       3198 AAAGGCTGATCCTTCCAACGCCGGACACGTAGCAGCACCATTCGCAGGCGTTGTCACAGT 3257

Qy       3453 GACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTAT 3512
              |||  | |  ||||| | |   ||||| || || ||| |||| ||  |||| ||||| ||
Db       3258 GACAATCGAGGAAGGCGCTACCGTCAAAGCCGGCGATCCAGTTGCTGTCATTGAGGCAAT 3317

Qy       3513 GAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCC 3572
              ||||||||| ||||| ||| ||||  |   ||||| |   |||| ||  | ||  |  |
Db       3318 GAAGATGGAGGCAACGATCTCTGCCACCACAGACGGAACCGTCGACCGTATCGTCTTGAC 3377

Qy       3573 TGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGT 3611
                 || || ||||| |||||||| |||||| || ||||
Db       3378 GCAGGCCACCAAGGTCGAAGGTGGAGACTTGCTCCTCGT 3416
```

RESULT 11
AAF71419
ID AAF71419 standard; DNA; 1719 BP.
XX
AC AAF71419;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:119.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000943.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-061975/07.
DR P-PSDB; AAB79302.
XX
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes.
XX
PS Claim 3; Page 316-319; 1246pp; English.
XX
CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for
CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteinogenic or
CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein regions required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)
XX
SQ Sequence 1719 BP; 405 A; 457 C; 468 G; 389 T; 0 U; 0 Other;

Query Match 47.2%; Score 1708; DB 4; Length 1719;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1719; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy         48 GAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGAT 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGTCGTGCAGGTCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGAT 60

Qy        108 GTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGG 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTTGATTGGGGGAATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGG 120

Qy        168 CGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCC 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCC 180

Qy        228 AGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGC 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGC 240

Qy        288 AGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATT 347
```

```
Db        241 AGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATT 300
Qy        348 CCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGC 407
Db        301 CCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGC 360
Qy        408 GTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTA 467
Db        361 GTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTA 420
Qy        468 CCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGG 527
Db        421 CCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGG 480
Qy        528 CATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC 587
Db        481 CATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC 540
Qy        588 GGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAA 647
Db        541 GGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAA 600
Qy        648 CATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGT 707
Db        601 CATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGT 660
Qy        708 TGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATT 767
Db        661 TGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATT 720
Qy        768 AGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGA 827
Db        721 AGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGA 780
Qy        828 ACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGA 887
Db        781 ACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGA 840
Qy        888 AGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGA 947
Db        841 AGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGA 900
Qy        948 AATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGT 1007
Db        901 AATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGT 960
Qy       1008 AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGA 1067
Db        961 AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGA 1020
Qy       1068 AAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGAC 1127
Db       1021 AAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGAC 1080
Qy       1128 TGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAAC 1187
Db       1081 TGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAAC 1140
Qy       1188 CTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTG 1247
Db       1141 CTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTG 1200
Qy       1248 CCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGC 1307
Db       1201 CCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGC 1260
Qy       1308 GTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGA 1367
Db       1261 GTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGA 1320
Qy       1368 AATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGA 1427
Db       1321 AATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGA 1380
Qy       1428 AACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAAC 1487
Db       1381 AACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAAC 1440
Qy       1488 CAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGC 1547
Db       1441 CAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGC 1500
Qy       1548 CACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCA 1607
Db       1501 CACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCA 1560
Qy       1608 GGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCC 1667
Db       1561 GGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCC 1620
Qy       1668 AAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACG 1727
Db       1621 AAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACG 1680
Qy       1728 CGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
Db       1681 CGGTTCCCGTGACCG-CTGAAGCAGCTTGGCCCAGCCGCG 1719
```

RESULT 12
AAF71420
ID AAF71420 standard; DNA; 1406 BP.
XX
AC AAF71420;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:121.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000943.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.

PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-061975/07.
DR P-PSDB; AAB79303.
XX
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes.
XX
PS Claim 3; Page 321-323; 1246pp; English.
XX
CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for
CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteinogenic or
CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein regions required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH)
XX
SQ Sequence 1406 BP; 325 A; 381 C; 385 G; 315 T; 0 U; 0 Other;

Query Match 38.5%; Score 1395; DB 4; Length 1406;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1406; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 60

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 120

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 180

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 240

Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 300

Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 360

Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 420

Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 480

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 540

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 600

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 660

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 720

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 780

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 840

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 900

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 960

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1020

Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1080

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1140

Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1200

Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1260

Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1320

Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1380

Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
              || |||||||||||||||||||||||||
Db       1381 CG-CTGAAGCAGCTTGGCCCAGCCGCG 1406
```

RESULT 13
ACA37896
ID ACA37896 standard; DNA; 3381 BP.
XX
AC ACA37896;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #19553.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycobacterium avium.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU34026.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 25766; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 3381 BP; 542 A; 1154 C; 1196 G; 489 T; 0 U; 0 Other;

Query Match 37.6%; Score 1361.6; DB 7; Length 3381;
Best Local Similarity 63.2%; Pred. No. 0;
Matches 2131; Conservative 0; Mismatches 1234; Indels 9; Gaps 2;

```
Qy   238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db    10 AAAGTCCTGGTCGCCAACCGCGGGGAGATCGCGATCCGCGCGTTCCGCGCGGCCTACGAG 69

Qy   298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db    70 CTGGAAATGGCCACCGTGGCGGTGTATCCCTACGAGGACCGCAATTCGGTGCACCGGTTG 129

Qy   358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db   130 AAGGCCGACGAGTCCTACCAGATCGGTGAGGAGGGCCACCCAGTCCGCGCCTACCTGTCC 189

Qy   418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db   190 GTCGACGAGATCGTCGGCACGGCGCTGGCCTGCGGCGCGGATGCGATCTACCCCGGCTAC 249

Qy   478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db   250 GGCTTCCTGTCGGAGAACCCCGATCTGGCGGCGGCGTGCGCGGCGGCCGGCATCACGTTC 309

Qy   538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db   310 GTCGGCCCCAGCGCCGAGGTGCTCGAGCTCACCGGCGACAAGTCGCGGGCCATCGCGGCG 369

Qy   598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db   370 GCGCGGGCCGCGGGTCTTCCGGTGCTGGCCTCCTCGCCGCCGTCGACCTCGGTGCAGGAG 429

Qy   658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db   430 TTGCTGAGCGCCGCCGAGACGATGACGTTCCCGCTGTTCGTCAAGGCGGTCGCCGGCGGC 489

Qy   718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db   490 GGCGGCCGGGGGATGCGCCGGGTCACCGATCCGGGGGCGCTGGCCGAGGCGATCGAGGCG 549

Qy   778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db   550 GCCAGCCGTGAGGCCGAGTCCGCGTTCGGGGACGCCTCGGTGTTTCTCGAGCAGGCGGTG 609

Qy   838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db   610 ATCAATCCGCGGCACATCGAGGTCCAGATCCTGGCCGACACGCACGGCAACGTGATGCAC 669

Qy   898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db   670 CTCTACGAGCGGGACTGCAGCGTGCAGCGCCGGCACCAGAAGGTGATCGAGATCGCCCCG 729

Qy   958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db   730 GCGCCCAACCTGGATCCGGCGCTGCGCGAACGTATCTGCGCCGACGCGGTGGCGTTCGCA 789

Qy  1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db   790 CGCAGCATCGGCTACACCTGCGCGGGCACTGTGGAGTTCCTGCTCGACGAACGCGGCAAC 849

Qy  1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db   850 CACGTGTTCATCGAGATGAACCCGCGCATCCAGGTGGAGCACACGGTGACCGAGGAGATC 909

Qy  1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db   910 ACCGACGTCGACCTGGTGTCCGCGCAGCTGCGCATCGCCTCCGGCCAGACGCTGGAAGAG 969

Qy  1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
```

```
Db        970 ATTGGCCTGAGCCAGGATTCGGTCATCTCACGCGGCGCGGCCCTGCAGTGCCGGATCACC 1029
Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db       1030 ACCGAGGACCCCGCCAACGGTTTCCGCCCGGACACCGGCCGCATCACCGCCTACCGCACC 1089
Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db       1090 CCGGGTGGCGCGGGCATCCGGCTGGACGGCGGCACCACGCTGGGCGCGGAGATCAGCGCG 1149
Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db       1150 CACTTCGATTCGATGCTGATCAAGCTGACCTGCCGGGGACGCGACTTCCCCACGGCGGTG 1209
Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db       1210 CGGCGGGCGCGGCGCGCGGTCGCCGAATTCCGCATCCGCGGCGTCTCGACGAATATCCCG 1269
Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db       1270 TTCCTGCAAGCGGTTCTGGACGACCCGGATTTTCAGGCTGGGCGCATCACCACCTCGTTC 1329
Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db       1330 ATCGAACAGCGGCCGCAGTTGCTCACCGCGCGCAGTTCGGCCGACCGCGGCACCAAGATC 1389
Qy       1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db       1390 CTCAACTACCTGGCCGACGTGACCGTCAACAAGCCGCACGGCGAGCGCCCGTCGGCGGTG 1449
Qy       1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db       1450 TACCCGCATGACAAGCTGCCCGACATCGACCTGGCGGCGAGCCCGCCGGCGGGCTCCAAG 1509
Qy       1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db       1510 CAGCGGTTGACCGAACTCGGGCCGGAGCGATTCGCGGCCTGGCTGCGCGAGTCGGGCGGC 1569
Qy       1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db       1570 GTCGGCGTCACCGACACCACGTTCCGCGACGCGCACCAGTCGCTGCTGGCGACCCGGGTC 1629
Qy       1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db       1630 CGCACCAGCGGGCTGCTCAAGGTGGCACCCTACATCGCACGGACCATGCCGCAGCTGTTG 1689
Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db       1690 TCGGTGGAGTGTTGGGGCGGAGCGACTTACGACGTGGCGCTGCGGTTTCTCAAGGAGGAC 1749
Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db       1750 CCGTGGGAGCGGCTGACGGCGCTGCGCGAGGCGATGCCGAACATCTGCCTTCAGATGCTG 1809
Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db       1810 CTGCGTGGGCGTAACACCGTGGGGTACACGCCCTATCCGGAGACGGTCACGACGGCCTTC 1869
Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db       1870 GTCGCGGAGGCGACCGAGACCGGCATCGACATCTTCCGCATCTTCGACGCGTTGAACAAC 1929
Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db       1930 GTCGACTCGATGCGCCCGGCCATCGACGCGGTCCGCGAAACCGGAAGCGCCATAGCCGAA 1989
Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db       1990 GTGGCGATGTCCTACACCGGCGACCTGTCCGATCCGGCCGAAAAGCTTTACACGCTGGAC 2049
Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db       2050 TATTACCTGAAGCTGGCCGACCAGATCGTCGAAGCGGGCGCGCACGTGCTGGCGATCAAG 2109
```

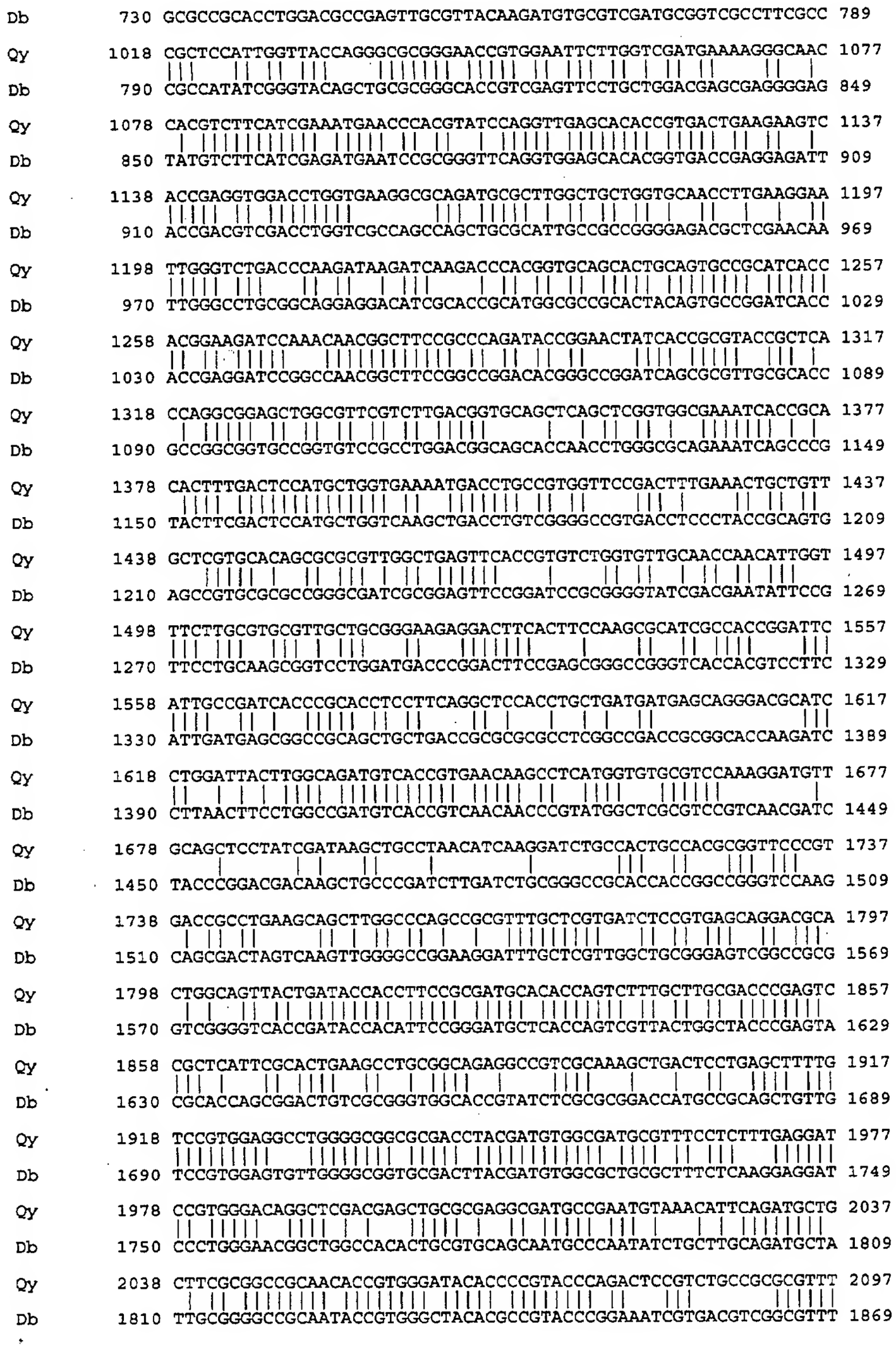
```
Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db       2110 GACATGGCCGGCCTGCTGCGCGCACCGGCGGCCGCGACGCTGGTGTCGGCGCTGAAGTCG 2169
Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db       2170 CGGTTCGACCTGCCGGTGCACGTGCACACCCATGACACGCCGGGCGGACAGCTGGCGACC 2229
Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db       2230 TACGTGGCGGCCTGGCAGGCCGGGGCCGATGCGGTCGACGGCGCCGCCGCGCCGCTCGCC 2289
Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db       2290 GGAACCACCAGCCAGCCGGCACTATCGTCGATCGTGGCCGCCGCGGCCAACACGGAGTAT 2349
Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db       2350 GACACCGGGTTGTCCCTGCCGGCGGTGTGCGATCTGGAGCCGTACTGGGAGGCGCTGCGA 2409
Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db       2410 AAGGTATACGCTCCCTTCGAATCCGGTCTTCCGGCGCCGACCGGGCGGGTGTATCACCAC 2469
Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db       2470 GAGATCCCGGGCGGCCAATTGTCGAATCTGCGTCAGCAGGCGATTGCGCTGGGACTCGGC 2529
Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db       2530 GACCGGTTCGAGGACATCGAAAACGCCTACGCCGGCGCCGACGCGATTTTGGGACACCTG 2589
Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db       2590 GTCAAGGTCACACCGTCCAGCAAGGTGGTGGGCGACCTGGCGCTGGCCCTGGTGGGTGCG 2649
Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db       2650 GGTGTGGGCGCGCAGGACTTCGCCGAAGACCCCTCCCGCTACGACATCCCGGATTCGGTG 2709
Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db       2710 ATCGGCTTTCTGCGCGGCGAGCTCGGTGACCCGCCGGGCGGCTGGCCGGAGCCGTTGCGC 2769
Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db       2770 ACCAAGGCCTTACAGGGCCGGGGGCCGGCCAAGCCGGAGCAGCCGCTGACGGCCGAGGAC 2829
Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db       2830 GAAGCGGCGTTGGCCGCCCCGGGTG------CGCGACGGCAGGCAGCGCTGAACCGGCTG 2883
Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db       2884 CTGTTCCCGGGACCGACCAAGGAGCTCGAGGAGCACCGCGAGCAGTACGGTGACACCTCG 2943
Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db       2944 GGGCTGAGCGCCAACCAGTTCTTCTACGGATTGCGGCAGGGCGACGAACACCGCGTCGAA 3003
Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db       3004 CT---CGAGCGCGGCGTGGAGTTGCTGATCGGCCTGGAGGCCATCTCCGATCCCGACGAA 3060
Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db       3061 CGCGGCATGCGCACCGTGATGTGCATTCTCAACGGGCAGCTGCGGCCGGTGGTGGTGCGC 3120
Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db       3121 GACCGCAGCATCGCCATCGACGTGCCCGCCGCCGAGAAGGCCGACCGCGCCAACCCGGAC 3180
```

```
Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
             ||  | |||||  || |||||  || ||  ||  ||  ||  |  |  |  |  ||      |||||
Db       3181 CACATCGCTGCCCCGTTCGCCGGGGTGGTGACGGTCGCCGCCGAGGTCGGCCGGCAGGTC 3240

Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
              ||||  ||   |  |  ||||  |  |||||||| ||||||||||||||   |  ||||
Db       3241 GAGGCCGGCCAGACCATCGCGACGATCGAGGCGATGAAGATGGAAGCCGCGGTCACCAGC 3300

Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
              |        |||||  | |  ||  | |  ||  |      |  |  | || ||  || |||
Db       3301 CCGAAGTCGGGCAAGGTGGCGCGAATCGCCGTGTCGCGGACCGCTCAAGTCGAGGGCGGC 3360

Qy       3598 GACTTGATCGTCGT 3611
              |||  |  |  | ||
Db       3361 GACCTCCTGATGGT 3374
```

RESULT 14
ACA40756
ID ACA40756 standard; DNA; 3384 BP.
XX
AC ACA40756;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #22413.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU36886.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 28626; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3384 BP; 574 A; 1061 C; 1148 G; 601 T; 0 U; 0 Other;

Query Match 35.7%; Score 1294.4; DB 7; Length 3384;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
              |||  |  | ||  ||  ||  |||||  ||  ||||||  ||||  ||  ||  |||||  ||    ||||
Db         10 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 69

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
                 ||  |   |  ||  ||  ||   ||||  ||     ||  ||  ||     ||     |||||
Db         70 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 129

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
                 ||     ||  ||  ||  ||  ||     |   ||      ||  ||   |  ||  ||||||
Db        130 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 189

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
                 ||||  ||  ||   |||   |  ||         |      ||  ||  ||  ||  |||
Db        190 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 249

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
              ||  ||  ||  ||  ||  |||  |   |  ||  ||      |   ||  ||||      |||||  |   ||
Db        250 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 309

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
                 |  ||  ||  |  |  |  ||  ||  |||||  ||    |  ||   ||||||||||||    |   |  ||
Db        310 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 369

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
                 ||      |  ||  ||   ||||  ||   ||    |   ||   |  |||              ||||  ||
Db        370 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 429

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
                 |  |     ||  |   |||   |  |  |   ||  |  ||  ||  ||||||||||||||||||
Db        430 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 489

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
                 ||  ||  ||  |||||||||   |||  |          |  |  |||||   |            ||
Db        490 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 549

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
                 ||      ||  |||||  ||   ||||  |||||  ||      ||||  |||  |||   |  |||
Db        550 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 609

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
                 ||  ||  ||  |   ||  ||  ||  ||||||||  ||  |   ||   ||   ||  ||  ||   |  ||
Db        610 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 669

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
                 ||  |||||  ||||||||||    ||||||||||  ||  ||  ||  ||||   |  |||||
Db        670 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 729

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
                 ||  |  |||   ||||   |  ||   |||||  |      ||  ||  |   |||||  ||      |||   |
```


Db 730 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 789
Qy 1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db 790 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 849
Qy 1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db 850 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 909
Qy 1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db 910 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 969
Qy 1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db 970 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 1029
Qy 1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db 1030 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 1089
Qy 1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db 1090 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 1149
Qy 1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db 1150 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 1209
Qy 1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db 1210 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 1269
Qy 1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db 1270 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 1329
Qy 1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db 1330 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 1389
Qy 1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db 1390 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 1449
Qy 1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db 1450 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 1509
Qy 1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db 1510 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 1569
Qy 1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db 1570 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 1629
Qy 1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db 1630 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 1689
Qy 1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db 1690 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 1749
Qy 1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db 1750 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 1809
Qy 2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db 1810 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 1869
Qy 2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db 1870 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 1929
Qy 2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db 1930 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 1989
Qy 2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db 1990 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 2049
Qy 2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db 2050 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 2109
Qy 2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db 2110 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 2169
Qy 2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db 2170 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 2229
Qy 2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db 2230 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCCGCCGCTGGCG 2289
Qy 2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db 2290 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 2349
Qy 2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db 2350 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 2409
Qy 2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db 2410 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 2469
Qy 2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db 2470 GAGATTCCGGGCGGCCAACTGTCCAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 2529
Qy 2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db 2530 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 2589
Qy 2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db 2590 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 2649
Qy 2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db 2650 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 2709
Qy 2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db 2710 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 2769
Qy 2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db 2770 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 2826
Qy 3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db 2827 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 2883
Qy 3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db 2884 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 2943

```
Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db       2944 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 3003

Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db       3004 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 3060

Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db       3061 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 3120

Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db       3121 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 3180

Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
Db       3181 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTGCGTCGGCGAGCGGGTC 3240

Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
Db       3241 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 3300

Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
Db       3301 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 3360

Qy       3598 GACTTGATCGTCGTCGTTTCCTAA 3621
Db       3361 GACCTGTTGGTGGTGGTGAGCTGA 3384
```

RESULT 15
AAI99682_33/c
Continuation (34 of 45) of AAI99682 from base 3300001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682

	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000
WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 35.7%; Score 1294.4; DB 4; Length 110000;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db      23034 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 22975

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db      22974 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 22915

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db      22914 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 22855

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db      22854 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 22795

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db      22794 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 22735

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db      22734 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 22675

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db      22674 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 22615

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db      22614 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 22555

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db      22554 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 22495

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db      22494 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 22435

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db      22434 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 22375

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db      22374 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 22315

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db      22314 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 22255

Qy       1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db      22254 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 22195

Qy       1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
```

```
Db      22194 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 22135
Qy       1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db      22134 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 22075
Qy       1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db      22074 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 22015
Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db      22014 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 21955
Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db      21954 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 21895
Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db      21894 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 21835
Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db      21834 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 21775
Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db      21774 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 21715
Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db      21714 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 21655
Qy       1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db      21654 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 21595
Qy       1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db      21594 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 21535
Qy       1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db      21534 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 21475
Qy       1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db      21474 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 21415
Qy       1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db      21414 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 21355
Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db      21354 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 21295
Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db      21294 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 21235
Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db      21234 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 21175
Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db      21174 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 21115
Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db      21114 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 21055
```

```
Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db      21054 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 20995
Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db      20994 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 20935
Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db      20934 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 20875
Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db      20874 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 20815
Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db      20814 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCGCCGCTGGCG 20755
Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db      20754 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 20695
Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db      20694 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 20635
Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db      20634 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 20575
Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db      20574 GAGATTCCGGGCGGCCAACTGTCCAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 20515
Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db      20514 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 20455
Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db      20454 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 20395
Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db      20394 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 20335
Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db      20334 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 20275
Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db      20274 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 20218
Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db      20217 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 20161
Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db      20160 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 20101
Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db      20100 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 20041
Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db      20040 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 19984
```

```
Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
              || ||||| |  ||  ||  || | ||||||| ||| | || ||  |||  |||||
Db      19983 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 19924

Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
              |||||| | |||    || |    |  |||| || ||||| ||   |  |||    ||
Db      19923 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 19864

Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
              ||  | || || ||||| || || || ||||| ||     ||     |  || ||   ||||
Db      19863 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTGCGTCGGCGAGCGGGTC 19804

Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
              || ||  |  |  |||| | ||||||||| ||||||||||||||  | ||||| ||
Db      19803 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 19744

Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
              | |||| |||||   | || || || | |||  | |   |  |  |||||||  || ||
Db      19743 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 19684

Qy       3598 GACTTGATCGTCGTCGTTTCCTAA 3621
              ||| || | || || ||   || |
Db      19683 GACCTGTTGGTGGTGGTGAGCTGA 19660
```

Search completed: March 23, 2004, 18:35:43
Job time : 1318 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 17:20:33 ; Search time 237 Seconds
(without alignments)
8478.807 Million cell updates/sec

Title: US-10-045-072-1
Perfect score: 3621
Sequence: 1 tggggcggggttagatcctg.....tgatcgtcgtcgtttcctaa 3621

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		3621	100.0	3621	3	US-09-220-081-1	Sequence 1, Appli
2		3621	100.0	3621	4	US-09-677-575-1	Sequence 1, Appli
3	c	1294.4	35.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli
4	c	1294.4	35.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
5		1249.6	34.5	32155	4	US-08-311-731A-1	Sequence 1, Appli
6		509.2	14.1	3492	4	US-09-134-000C-2302	Sequence 2302, Ap
7		486.2	13.4	3465	4	US-09-134-001C-591	Sequence 591, App
8	c	479.8	13.3	5030	4	US-08-956-171E-324	Sequence 324, App
9		476.2	13.2	2922	4	US-09-107-532A-1156	Sequence 1156, Ap
10		313.4	8.7	1362	4	US-09-543-681A-1699	Sequence 1699, Ap
11		313	8.6	1362	1	US-07-956-700B-5	Sequence 5, Appli
12		313	8.6	1362	1	US-08-476-537-5	Sequence 5, Appli
13		313	8.6	1362	1	US-08-485-607-5	Sequence 5, Appli
14		313	8.6	1362	2	US-08-475-879-5	Sequence 5, Appli
15		313	8.6	1362	4	US-09-433-043B-5	Sequence 5, Appli
16		311.4	8.6	1362	1	US-08-611-107-7	Sequence 7, Appli
17		311.4	8.6	1362	2	US-08-422-560A-7	Sequence 7, Appli
18		311.4	8.6	1362	3	US-08-468-793-7	Sequence 7, Appli
19		299	8.3	3077	1	US-08-074-121-1	Sequence 1, Appli
20		299	8.3	3077	5	PCT-US94-06447-1	Sequence 1, Appli
21		298.6	8.2	3065	1	US-07-956-700B-1	Sequence 1, Appli
22		298.6	8.2	3065	1	US-08-476-537-1	Sequence 1, Appli
23		298.6	8.2	3065	1	US-08-485-607-1	Sequence 1, Appli
24		298.6	8.2	3065	1	US-08-611-107-5	Sequence 5, Appli
25		298.6	8.2	3065	2	US-08-422-560A-5	Sequence 5, Appli
26		298.6	8.2	3065	2	US-08-475-879-1	Sequence 1, Appli
27		298.6	8.2	3065	3	US-08-468-793-5	Sequence 5, Appli
28		298.6	8.2	3065	4	US-09-433-043B-1	Sequence 1, Appli
29	c	297.4	8.2	2907	4	US-09-252-991A-2354	Sequence 2354, Ap
30		288.6	8.0	1511	4	US-09-634-238-58	Sequence 58, Appl
31		284.8	7.9	1425	4	US-09-328-352-3436	Sequence 3436, Ap
32		270.6	7.5	1921	3	US-08-662-344-1	Sequence 1, Appli
33		263.8	7.3	1953	4	US-09-252-991A-2481	Sequence 2481, Ap
34		255.4	7.1	1664976	4	US-08-916-421B-1	Sequence 1, Appli
35		251.6	6.9	1356	4	US-09-540-236-1127	Sequence 1127, Ap
36	c	251.6	6.9	23673	4	US-09-773-816-1	Sequence 1, Appli
37	c	251.6	6.9	119211	4	US-09-596-002-40	Sequence 40, Appl
38		250.2	6.9	4403765	3	US-09-103-840A-2	Sequence 2, Appli
39		250.2	6.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
40		249	6.9	2031	4	US-09-252-991A-9572	Sequence 9572, Ap
41	c	249	6.9	3006	4	US-09-252-991A-9720	Sequence 9720, Ap
42	c	247.4	6.8	19702	4	US-08-961-527-7	Sequence 7, Appli
43		247	6.8	1398	4	US-09-252-991A-10409	Sequence 10409, A
44	c	247	6.8	1446	4	US-09-252-991A-10752	Sequence 10752, A
45		246.8	6.8	1830121	4	US-09-557-884-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-220-081-1
; Sequence 1, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-09-220-081-1

Query Match 100.0%; Score 3621; DB 3; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Qy        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Db        601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Qy        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Db        661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Qy        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Db        721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Qy        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Db        781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Db        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Db        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Db        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Db       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Db       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Db       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Db       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Db       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Db       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Qy       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Db       1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440

Qy       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Db       1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Qy       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Db       1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Qy       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Db       1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Qy       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Db       1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Qy       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Db       1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Qy       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db       1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Qy       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db       1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Qy       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db       1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db       1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db       1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db       2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db       2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db       2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db       2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
```

```
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580

Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640

Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700

Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760

Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820

Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880

Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940

Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060

Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120

Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180

Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240

Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300

Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360

Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420

Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480

Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540

Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600

Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 2
US-09-677-575-1
; Sequence 1, Application US/09677575
; Patent No. 6403351
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/677,575
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-09-677-575-1

Query Match 100.0%; Score 3621; DB 4; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120

Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180

Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240

Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300

Qy        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360

Qy        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420

Qy        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480

Qy        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540

Qy        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
```

```
Qy    601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Db    601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Qy    661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Db    661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Qy    721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Db    721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Qy    781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Db    781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Qy    841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Db    841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Qy    901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Db    901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Qy    961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Db    961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Qy   1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Db   1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Qy   1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Db   1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Qy   1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Db   1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Qy   1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Db   1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Qy   1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Db   1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Qy   1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Db   1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Qy   1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Db   1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Qy   1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Db   1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Qy   1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Db   1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Qy   1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Db   1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Qy   1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Db   1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
```

```
Qy   1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Db   1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Qy   1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db   1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Qy   1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db   1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Qy   1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db   1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy   1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db   1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy   1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db   1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy   2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db   2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy   2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db   2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy   2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db   2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy   2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db   2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy   2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db   2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy   2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db   2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy   2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db   2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy   2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db   2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy   2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db   2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy   2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db   2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy   2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db   2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy   2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db   2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy   2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
```

```
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 3
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 35.7%; Score 1294.4; DB 3; Length 4403765;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db    3317308 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 3317249
Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db    3317248 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 3317189
Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db    3317188 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 3317129
Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db    3317128 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 3317069
Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db    3317068 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 3317009
Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db    3317008 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 3316949
Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db    3316948 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 3316889
Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db    3316888 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 3316829
Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db    3316828 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 3316769
Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db    3316768 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 3316709
Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db    3316708 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 3316649
Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db    3316648 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 3316589
Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db    3316588 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 3316529
Qy       1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db    3316528 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 3316469
```

```
Qy       1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db    3316468 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 3316409

Qy       1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db    3316408 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 3316349

Qy       1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db    3316348 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 3316289

Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db    3316288 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 3316229

Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db    3316228 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 3316169

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db    3316168 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 3316109

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db    3316108 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 3316049

Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db    3316048 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 3315989

Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db    3315988 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 3315929

Qy       1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db    3315928 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 3315869

Qy       1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db    3315868 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 3315809

Qy       1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db    3315808 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 3315749

Qy       1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db    3315748 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 3315689

Qy       1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db    3315688 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 3315629

Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db    3315628 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 3315569

Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db    3315568 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 3315509

Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db    3315508 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 3315449

Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db    3315448 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 3315389

Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
```

```
Db    3315388 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 3315329

Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db    3315328 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 3315269

Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db    3315268 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 3315209

Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db    3315208 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 3315149

Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db    3315148 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 3315089

Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db    3315088 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCGCCGCTGGCG 3315029

Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db    3315028 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 3314969

Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db    3314968 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 3314909

Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db    3314908 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 3314849

Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db    3314848 GAGATTCCGGGCGGCCAACTGTCCAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 3314789

Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db    3314788 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 3314729

Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db    3314728 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 3314669

Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db    3314668 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 3314609

Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db    3314608 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 3314549

Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db    3314548 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 3314492

Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db    3314491 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 3314435

Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db    3314434 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 3314375

Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db    3314374 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 3314315

Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
```

```
Db    3314314 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 3314258

Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357

Db    3314257 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 3314198

Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417

Db    3314197 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 3314138

Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477

Db    3314137 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTGCGTCGGCGAGCGGGTC 3314078

Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537

Db    3314077 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 3314018

Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597

Db    3314017 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 3313958

Qy       3598 GACTTGATCGTCGTCGTTTCCTAA 3621

Db    3313957 GACCTGTTGGTGGTGGTGAGCTGA 3313934
```

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 35.7%; Score 1294.4; DB 3; Length 4411529;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2099; Conservative 0; Mismatches 1276; Indels 9; Gaps 3;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297

Db    3323034 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 3322975

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357

Db    3322974 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 3322915

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417

Db    3322914 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 3322855

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477

Db    3322854 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 3322795

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537

Db    3322794 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 3322735

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597

Db    3322734 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 3322675

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657

Db    3322674 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 3322615

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717

Db    3322614 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 3322555

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777

Db    3322554 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 3322495

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837

Db    3322494 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 3322435

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897

Db    3322434 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 3322375

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957

Db    3322374 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 3322315

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017

Db    3322314 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 3322255

Qy       1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077

Db    3322254 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 3322195

Qy       1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137

Db    3322194 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 3322135

Qy       1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197

Db    3322134 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 3322075

Qy       1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257

Db    3322074 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 3322015

Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317

Db    3322014 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 3321955

Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377

Db    3321954 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 3321895

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437

Db    3321894 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 3321835

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497

Db    3321834 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 3321775

Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557

Db    3321774 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 3321715

Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617

Db    3321714 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 3321655
```

```
Qy      1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db   3321654 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 3321595

Qy      1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db   3321594 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 3321535

Qy      1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db   3321534 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 3321475

Qy      1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db   3321474 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 3321415

Qy      1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db   3321414 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 3321355

Qy      1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db   3321354 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 3321295

Qy      1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db   3321294 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 3321235

Qy      2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db   3321234 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 3321175

Qy      2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db   3321174 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 3321115

Qy      2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db   3321114 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 3321055

Qy      2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db   3321054 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 3320995

Qy      2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db   3320994 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 3320935

Qy      2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db   3320934 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 3320875

Qy      2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db   3320874 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 3320815

Qy      2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db   3320814 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCGCCGCTGGCG 3320755

Qy      2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db   3320754 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 3320695

Qy      2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db   3320694 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 3320635

Qy      2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db   3320634 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 3320575

Qy      2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
```

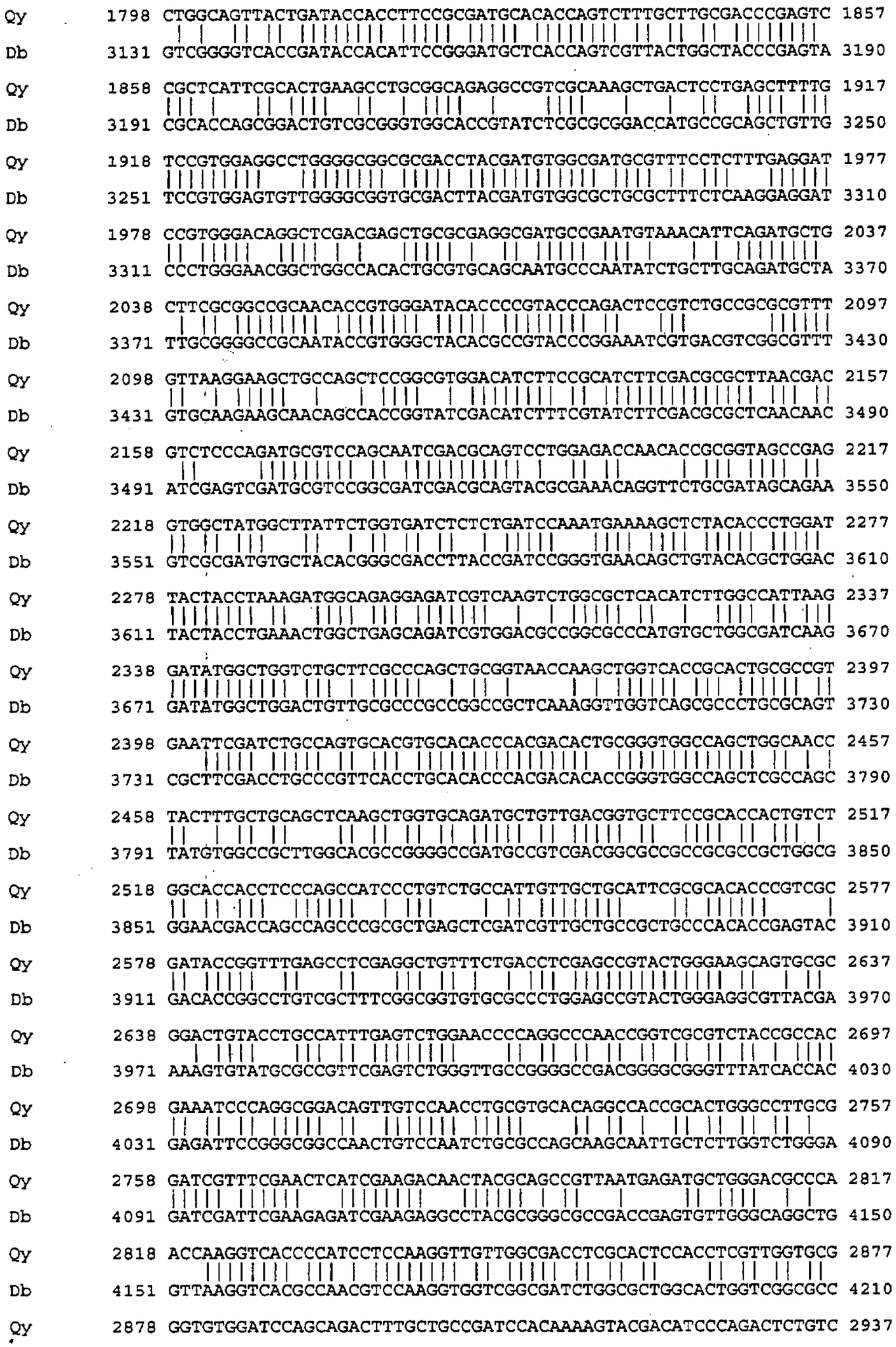
```
Db   3320574 GAGATTCCGGGCGGCCAACTGTCCAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 3320515

Qy      2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db   3320514 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 3320455

Qy      2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db   3320454 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 3320395

Qy      2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db   3320394 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 3320335

Qy      2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db   3320334 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 3320275

Qy      2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db   3320274 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 3320218

Qy      3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db   3320217 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 3320161

Qy      3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db   3320160 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 3320101

Qy      3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db   3320100 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 3320041

Qy      3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db   3320040 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 3319984

Qy      3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db   3319983 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 3319924

Qy      3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db   3319923 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 3319864

Qy      3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
Db   3319863 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTCGTCGGCGAGCGGGTC 3319804

Qy      3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
Db   3319803 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 3319744

Qy      3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
Db   3319743 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 3319684

Qy      3598 GACTTGATCGTCGTCGTTTCCTAA 3621
Db   3319683 GACCTGTTGGTGGTGGTGAGCTGA 3319660
```

RESULT 5
US-08-311-731A-1
; Sequence 1, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM TUBERCULOSIS
US-08-311-731A-1

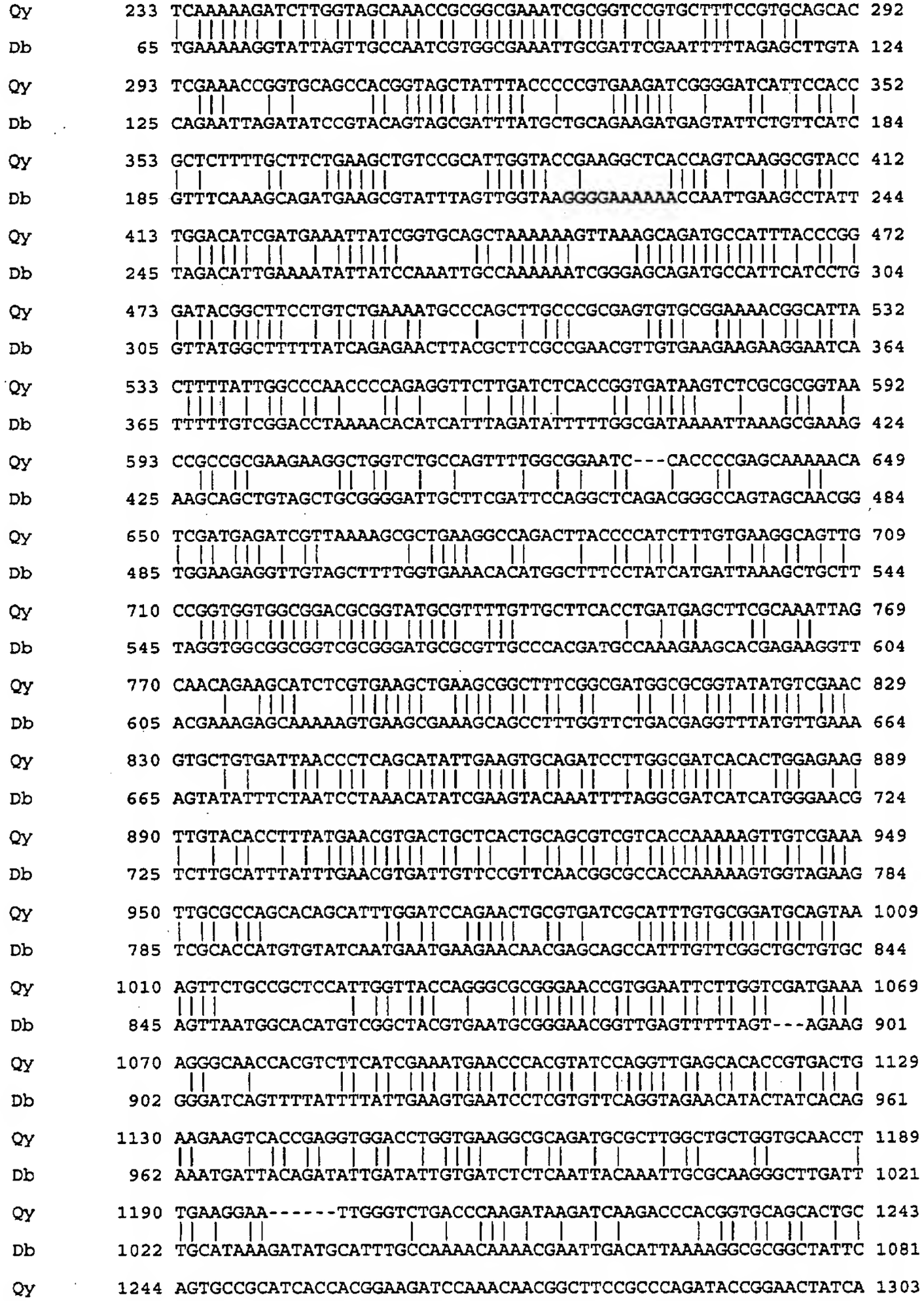
Query Match 34.5%; Score 1249.6; DB 4; Length 32155;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 1274; Indels 13; Gaps 7;

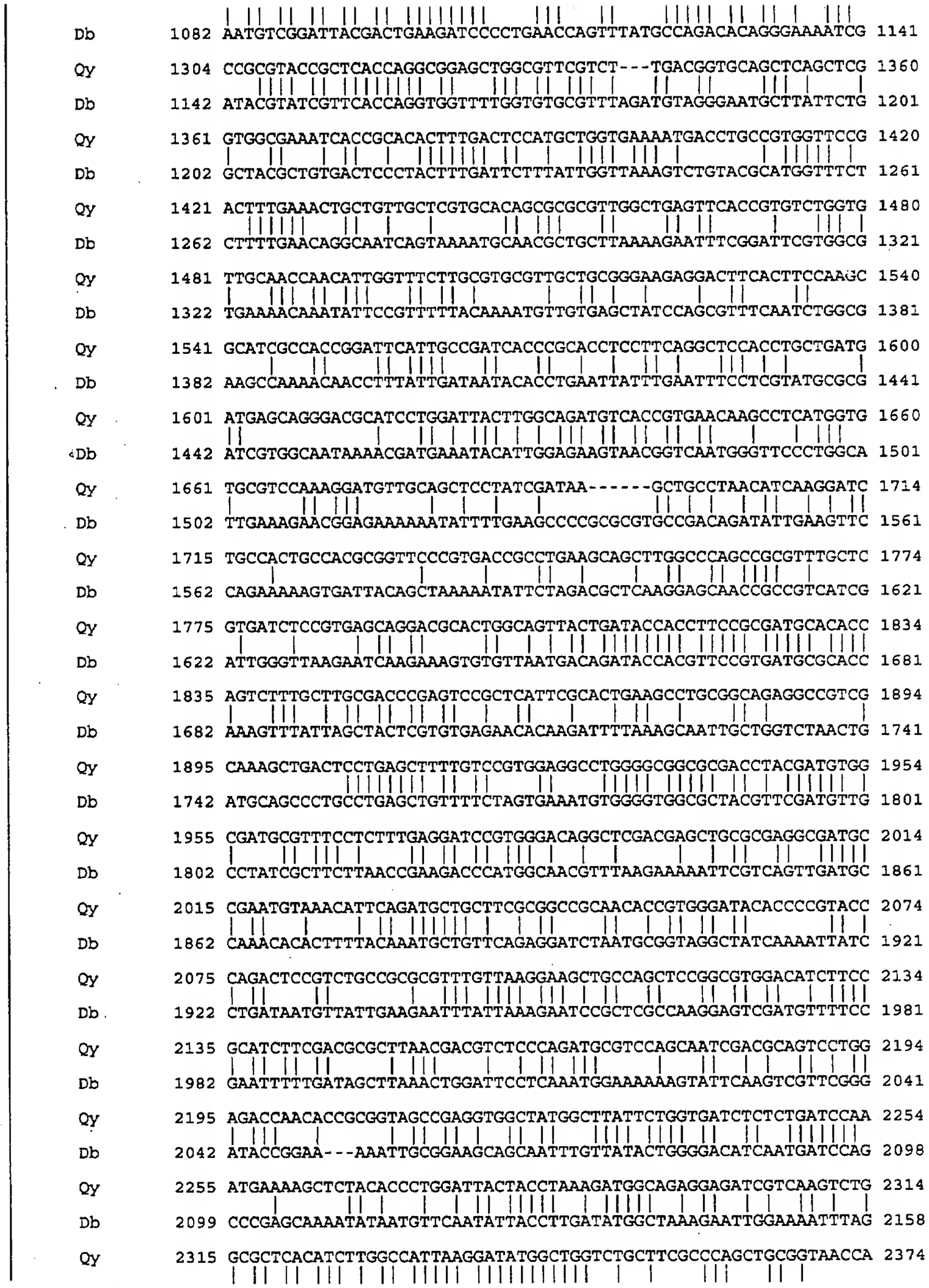
```
Qy   238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db  1574 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 1633
Qy   298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db  1634 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 1693
Qy   358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db  1694 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 1753
Qy   418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db  1754 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 1813
Qy   478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db  1814 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 1873
Qy   538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db  1874 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 1933
Qy   598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db  1934 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 1993
Qy   658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db  1994 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 2053

Qy   718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db  2054 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 2113
Qy   778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db  2114 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 2173
Qy   838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db  2174 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 2233
Qy   898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db  2234 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 2293
Qy   958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db  2294 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 2353
Qy  1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db  2354 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 2413
Qy  1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db  2414 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 2473
Qy  1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db  2474 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 2533
Qy  1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db  2534 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 2593
Qy  1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db  2594 ACCGAGGATCCGGCCAACGGCTTCCGGCC-GACACGGGCCGGATCAGCGCGTTGCGACCC 2652
Qy  1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db  2653 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCAC-CAACCTGGCGCAGAAATCAGCCCG 2711
Qy  1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db  2712 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCG-GGCCGTGACCTCCCTACCGCAGTG 2770
Qy  1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db  2771 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 2830
Qy  1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db  2831 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 2890
Qy  1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db  2891 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 2950
Qy  1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db  2951 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 3010
Qy  1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db  3011 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 3070
Qy  1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db  3071 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 3130
```



RESULT 6
US-09-134-000C-2302
; Sequence 2302, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2302
; LENGTH: 3492
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2302

Query Match 14.1%; Score 509.2; DB 4; Length 3492;
Best Local Similarity 49.1%; Pred. No. 1.1e-134;
Matches 1661; Conservative 0; Mismatches 1668; Indels 51; Gaps 10;



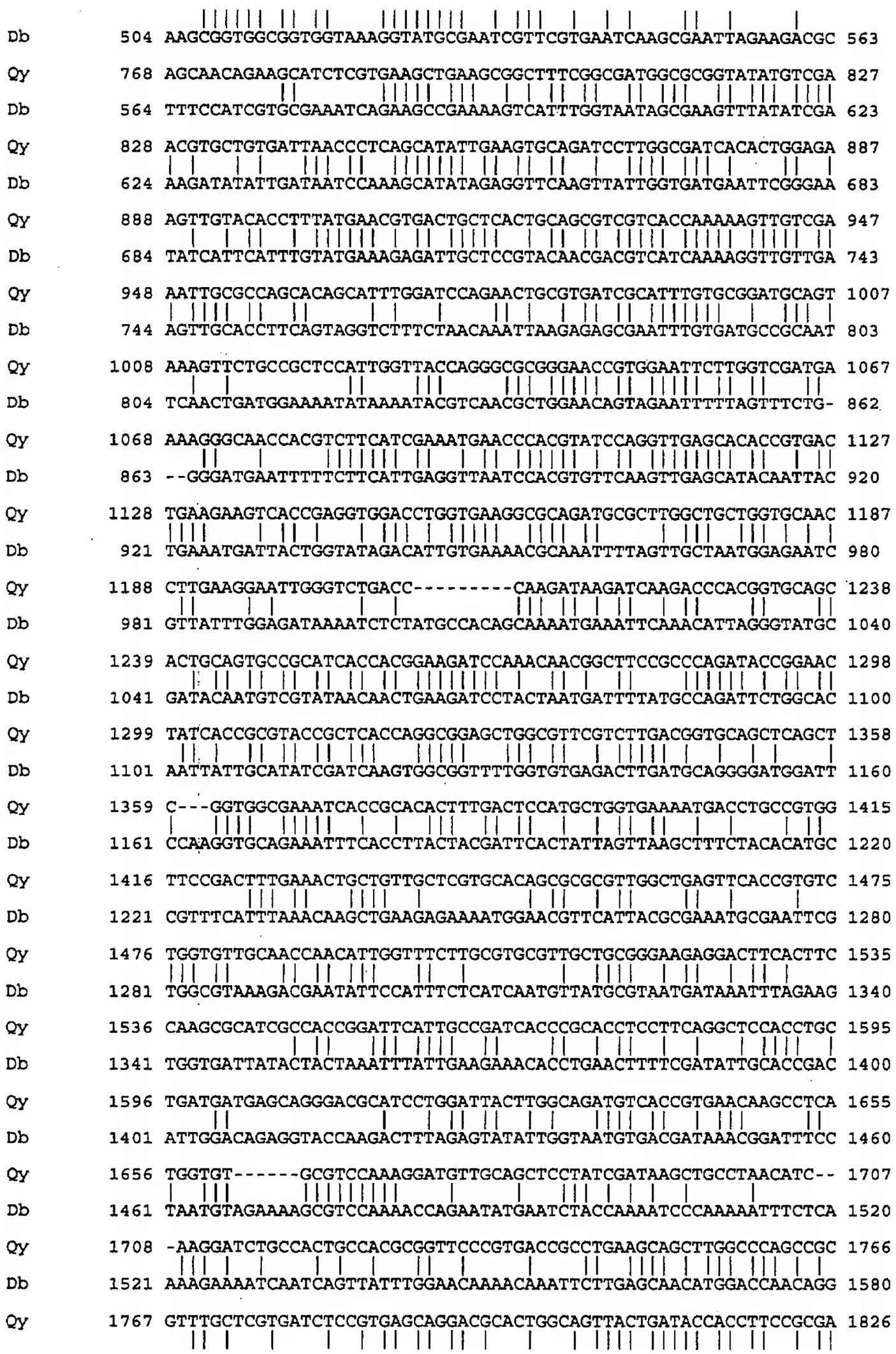


```
Db       2159 GTGCACAAATCATTGCGATTAAAGATATGGCTGGCTTATTGAAACCACAAGCTGCTTATC 2218
Qy       2375 AGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACA 2434
Db       2219 GTTTAATTAGTGAATTAAAGGCAGCCACGGATTTACCAATTCATCTCCACACTCACGACA 2278
Qy       2435 CTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTG 2494
Db       2279 CTAGTGGCAATGGGATCATCACATATTCAGCAGCTACTAAAGCGGGCGTTGATATTGTTG 2338
Qy       2495 ACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTG 2554
Db       2339 ACGTCGCAATGAGTGCGATGAGTGGCGCAACTAGCCAGCCAAGTATGAACAGTCTATATT 2398
Qy       2555 CTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCG 2614
Db       2399 ATGCTTTAGTCAATGGTGAACGGACGCCAACTATTAACATTGATAATGCACAGAAAATCA 2458
Qy       2615 AGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCC 2674
Db       2459 ATCATTATTGGGAAGATGTGCGCATGTATTATCAACCATTTGAAAATGGGTTAAATGCCC 2518
Qy       2675 CAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCAC 2734
Db       2519 CGCAAACAGAAGTCTATATGCATGAAATGCCTGGTGGGCAATACTCTAATCTACAGCAGC 2578
Qy       2735 AGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCG 2794
Db       2579 AAGCAAAAGCGGTAGGCTTAGGGCACCGTTGGGATGAAATCAAAAAAATATATCACACAG 2638
Qy       2795 TTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACC 2854
Db       2639 TGAATTTGATGTTTGGCGATATTGTAAAAGTGACGCCATCTTCAAAAGTCGTGGGAGACA 2698
Qy       2855 TCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAA 2914
Db       2699 TGGCACTCTTCATGGTTCAAAATAATCTGACAGAACAAGATGTTTATGCGCGTGGTGAAG 2758
Qy       2915 AGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAG 2974
Db       2759 AACTAAGTTTTCCTGAATCAGTTGTGACTTTTTTCCAAGGTGATTTAGGTCAACCAGTTG 2818
Qy       2975 GTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCAC 3034
Db       2819 GTGGGTTCCCGAAAGAGCTGCAACGAATTATTCTAAAAGGTCGACCAGCCTTCACAGAGC 2878
Qy       3035 CTCTGACGGA---AGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGG 3091
Db       2879 GGCCAGGTGACTTAGCAGCACCTGTTGATTTTGCGAAAGTACAAGAAGAATTAGCCGAAA 2938
Qy       3092 AACGTCGCAATAGCCTCAACCGCC----------------TGCTGTTCCCGAAGCCAA 3133
Db       2939 AAATTGGGTATCAACCTAAATTGGAAGAAGTTTTAAGTTATTTAATGTATCCACAAGTAT 2998
Qy       3134 CCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTG 3193
Db       2999 TTTTAGAGTATCGTCAAAAATACGAAACCTTTGGAGATATAACCTTATTAGATACGCCAA 3058
Qy       3194 AATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCA 3253
Db       3059 CATTCTTTAATGGTATTCGTCAAGGA---GAGACATTGGAAGTTCAAATTGAACGAGGAA 3115
Qy       3254 CCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATG 3313
Db       3116 AAACGTTAATTATTCGTTTAGATGAAATTGGCGAACCTGATATTGACGGAAACCGTGTCT 3175
Qy       3314 TTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGT 3373
Db       3176 TGTTCTTCAATTTGAATGGGCAACGTCGTGAGGTTTTAGTTAAAGATGCCTCTATTAAGT 3235
Qy       3374 CTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCAT 3433
Db       3236 CTGCTGTTCAAGTAAAACAAAAGGCGGAACCAACTAATAAAGAACAAATCGGCGCAACGA 3295

Qy       3434 TCGCTGG---TGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATG 3490
Db       3296 TGTCTGGTTCTGTTCTACAAGTATTGGTCAAACGTGGCGATAAAGTAGAAAAAGGTCAGC 3355
Qy       3491 CAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCA 3550
Db       3356 CGTTGCTGATTACAGAAGCCATGAAAATGGAAACGACTATCGAAGCACGTTTTGCGGGCA 3415
Qy       3551 AAATCGATCGCGTTGTGGTT 3570
Db       3416 CGGTAGATCACATTTATGTT 3435
```

RESULT 7
US-09-134-001C-591
; Sequence 591, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 591
; LENGTH: 3465
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-591

Query Match 13.4%; Score 486.2; DB 4; Length 3465;
Best Local Similarity 48.4%; Pred. No. 3.9e-128;
Matches 1661; Conservative 0; Mismatches 1718; Indels 54; Gaps 9;

```
Qy        231 ATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGC 290
Db         24 AATAAAGAAATTACTTGTTGCTAACCGTGGTGAAATCGCCATTAGAATTTTTAGAGCGGC 83
Qy        291 ACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCA 350
Db         84 AGCAGAATTAAATATCAGTACAGTAGCAATTTATTCTAATGAAGATAAAAGTTCGTTACA 143
Qy        351 CCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTA 410
Db        144 TAGATATAAAGCAGATGAATCCTATCTAGTTGGAAGTGATTTAGGACCTGCTGAAAGTTA 203
Qy        411 CCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCC 470
Db        204 TTTGAATATCGAACGTATCATCGAAGTAGCTCTTCGCGCAGGTGTCGATGCAATTCATCC 263
Qy        471 GGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCAT 530
Db        264 TGGGTATGGTTTTTTAAGTGAAAATGAACAATTTGCACGCCGATGTGCTGAGGAAGGCAT 323
Qy        531 TACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGT 590
Db        324 TAAATTTATAGGTCCGCATCTTGAACATCTAGACATGTTTGGAGATAAGGTTAAGGCTAG 383
Qy        591 AACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGG---AATCCACCCCGAGCAAAAA 647
Db        384 AACAACTGCTATTAACGCTAACTTACCTGTAATCCCGGGTACAGATGGTCCTATTGAAAG 443
Qy        648 CATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGT 707
Db        444 TTTTGAAGCTGCAGAACAGTTTGCTAATGAAGCAGGTTACCCACTTATGATTAAGGCCAC 503
Qy        708 TGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATT 767
```


Db 504 AAGCGGTGGCGGTGGTAAAGGTATGCGAATCGTTCGTGAATCAAGCGAATTAGAAGACGC 563
Qy 768 AGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGA 827
Db 564 TTTCCATCGTGCGAAATCAGAAGCCGAAAAGTCATTTGGTAATAGCGAAGTTTATATCGA 623
Qy 828 ACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGA 887
Db 624 AAGATATATTGATAATCCAAAGCATATAGAGGTTCAAGTTATTGGTGATGAATTCGGGAA 683
Qy 888 AGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGA 947
Db 684 TATCATTCATTTGTATGAAAGAGATTGCTCCGTACAACGACGTCATCAAAAGGTTGTTGA 743
Qy 948 AATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGT 1007
Db 744 AGTTGCACCTTCAGTAGGTCTTTCTAACAAATTAAGAGAGCGAATTTGTGATGCCGCAAT 803
Qy 1008 AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGA 1067
Db 804 TCAACTGATGGAAAATATAAAATACGTCAACGCTGGAACAGTAGAATTTTTAGTTTCTG- 862
Qy 1068 AAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGAC 1127
Db 863 --GGGATGAATTTTTCTTCATTGAGGTTAATCCACGTGTTCAAGTTGAGCATACAATTAC 920
Qy 1128 TGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAAC 1187
Db 921 TGAAATGATTACTGGTATAGACATTGTGAAAACGCAAATTTTAGTTGCTAATGGAGAATC 980
Qy 1188 CTTGAAGGAATTGGGTCTGACC----------CAAGATAAGATCAAGACCCACGGTGCAGC 1238
Db 981 GTTATTTGGAGATAAAATCTCTATGCCACAGCAAAATGAAATTCAAACATTAGGGTATGC 1040
Qy 1239 ACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAAC 1298
Db 1041 GATACAATGTCGTATAACAACTGAAGATCCTACTAATGATTTTATGCCAGATTCTGGCAC 1100
Qy 1299 TATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCT 1358
Db 1101 AATTATTGCATATCGATCAAGTGGCGGTTTTGGTGTGAGACTTGATGCAGGGGATGGATT 1160
Qy 1359 C---GGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGG 1415
Db 1161 CCAAGGTGCAGAAATTTCACCTTACTACGATTCACTATTAGTTAAGCTTTCTACACATGC 1220
Qy 1416 TTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTC 1475
Db 1221 CGTTTCATTTAAACAAGCTGAAGAGAAAATGGAACGTTCATTACGCGAAATGCGAATTCG 1280
Qy 1476 TGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTC 1535
Db 1281 TGGCGTAAAGACGAATATTCCATTTCTCATCAATGTTATGCGTAATGATAAATTTAGAAG 1340
Qy 1536 CAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGC 1595
Db 1341 TGGTGATTATACTACTAAATTTATTGAAGAAACACCTGAACTTTTCGATATTGCACCGAC 1400
Qy 1596 TGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCA 1655
Db 1401 ATTGGACAGAGGTACCAAGACTTTAGAGTATATTGGTAATGTGACGATAAACGGATTTCC 1460
Qy 1656 TGGTGT------GCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC-- 1707
Db 1461 TAATGTAGAAAAGCGTCCAAAACCAGAATATGAATCTACCAAAATCCCAAAAATTTCTCA 1520
Qy 1708 -AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGC 1766
Db 1521 AAAGAAAATCAATCAGTTATTTGGAACAAAACAAATTCTTGAGCAACATGGACCAACAGG 1580
Qy 1767 GTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGA 1826
Db 1581 TGTTACAAATTGGGTTAGAGAACAAGAAGATGTTTTAATTACCGATACTACATTTAGAGA 1640
Qy 1827 TGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGA 1886
Db 1641 TGCACACCAATCTTTACTTGCAACACGTGTAAGAACAAAAGATATGATGAACATTGCATC 1700
Qy 1887 GGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTA 1946
Db 1701 TAAAACTGCTGAAGTTTTTAAAGATAGTTTTTCATTAGAAATGTGGGGTGGTGCAACATT 1760
Qy 1947 CGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGA 2006
Db 1761 TGATGTCGCCTATAATTTCTTGAAAGAGAATCCATGGGAACGTTTAGAAAGATTGCGCAA 1820
Qy 2007 GGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACAC 2066
Db 1821 AGCCATTCCGAATGTGTTATTCCAAATGTTATTACGAGCTTCGAACGCAGTAGGTTATAA 1880
Qy 2067 CCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGA 2126
Db 1881 AAACTATCCTGATAATGTAATTAAGAAATTCGTTCATGAAAGTGCAAAAGCTGGTGTAGA 1940
Qy 2127 CATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGC 2186
Db 1941 TGTTTTCCGTATATTCGACTCATTGAACTGGGTTGATCAAATGAAAGTAGCGAATGAAGC 2000
Qy 2187 AGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTC 2246
Db 2001 TGTTCAAGAAGCTGGAATGGTATCTGAGGGTACAATTTGCTATACAGGTGATATTTTAAA 2060
Qy 2247 TGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGT 2306
Db 2061 TGCTGAACGTTCCAATATTTATACTTTAGATTATTACGTTAAAATGGCTAAAGAACTGGA 2120
Qy 2307 CAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGC 2366
Db 2121 AAGAGAAGGATTCCATATATTAGCAATTAAAGATATGGCTGGTTTATTGAAACCGAAAGC 2180
Qy 2367 GGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACAC 2426
Db 2181 AGCTTACGAATTAATTGGTGAATTACGTGAGGCAACACATCTTCCAATTCATTTACATAC 2240
Qy 2427 CCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGA 2486
Db 2241 ACATGATACTAGTGGAAATGGATTGTTGACATATAAACAAGCAATTGATGCTGGCGTAGA 2300
Qy 2487 TGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGC 2546
Db 2301 TATTATAGATACTGCTGTTGCATCTATGAGTGGTTTAACGAGTCAACCAAGTGCAAATTC 2360
Qy 2547 CATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTC 2606
Db 2361 ATTATATTATGCACTAAATGGATTTCCACGTAATTTAAGAACTGATATTGATGGGTTAGA 2420
Qy 2607 TGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAAC 2666
Db 2421 AGAGTTGAGTCATTACTGGTCTGTAGTCAGACCTTACTATGCAGACTTTGAGAGTGATAT 2480
Qy 2667 CCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCT 2726
Db 2481 CAAATCACCAAATACAGAAATTTATCAACATGAAATGCCAGGTGGCCAATATTCAAACTT 2540
Qy 2727 GCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTA 2786
Db 2541 AAGTCAACAAGCTAAAAGTTTAGGATTGGGCGAACGTTTTGATGAAGTCAAAGAGATGTA 2600
Qy 2787 CGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGT 2846
Db 2601 TCGTCGTGTCAACTTCCTGTTTGGAGATCTTGTAAAAGTAACACCATCTTCAAAGGTAGT 2660
Qy 2847 TGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGA 2906
Db 2661 TGGAGATATGGCACTATATATGGTGCAAAATGATCTTGATGAAGATACGGTCATCAATGA 2720

```
Qy       2907 TCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAA 2966
Db       2721 TGGTTATAAATTAGATTTCCCAGAATCTGTTGTGTCATTCTTTAAAGGTGACATTGGACA 2780

Qy       2967 CCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGG 3026
Db       2781 ACCTGTCAACGGATTCAACAAGAAATTGCAAGATGTTATTTTAAAAGGACAGCAACCAAT 2840

Qy       3027 CAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTC 3086
Db       2841 TACTGAAAGACCAGGTGAATACTTGGAGCCGGTCGATTTTGAAGCAATCCGTCAAGAATT 2900

Qy       3087 CAAGGA--------------------ACGTCGCAATAGCCTCAACCGCCTGCTGTTCCC 3125
Db       2901 AAGCGACATACAACAAGACGAGGTAACAGAACAAGATATAATTAGTTATGTACTTTATCC 2960

Qy       3126 GAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGA 3185
Db       2961 GAAGGTATATAAACAATATATTCAAACGAAAGAGCAATTTGGTAATGTATCTTTACTGGA 3020

Qy       3186 TGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGA 3245
Db       3021 TACACCGACATTCTTATTTGGCATGCGTAATG---GTGAAACAGTTGAAATTGAAATTGA 3077

Qy       3246 TGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTAT 3305
Db       3078 TACTGGTAAACGTCTAATTATTAAATTAGAAACAATCAGTGAACCAGATGAAAATGGTAA 3137

Qy       3306 GCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTC 3365
Db       3138 GCGTACAATTTATTACGCTATGAATGGTCAAGCAAGACGTATCTATATTCAAGATGAAAA 3197

Qy       3366 CGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGC 3425
Db       3198 TGTTAAAACGAATGCTAATGTTAAACCTAAGGCGGATAAATCAAATCCAAATCATATTGG 3257

Qy       3426 TGCACCATTCGCTGGTGTTGTCAC---CGTGACTGTTGCTGAAGGTGATGAGGTCAAGGC 3482
Db       3258 TGCTCAAATGCCTGGTTCTGTAACTGAAGTCAAAGTGTCTGTAGGCGATGAAGTTCAAGC 3317

Qy       3483 TGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGT 3542
Db       3318 TAATCAGCCATTATTAATCACTGAAGCAATGAAGATGGAAACGACGATTCAGGCACCATT 3377

Qy       3543 TGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTT 3602
Db       3378 TGATGGAATTATTAAACAAATCAATGTTGCTAATGGAGATGCCATTGCCACAGGAGATTT 3437

Qy       3603 GATCGTCGTCGTT 3615
Db       3438 ATTAGTGGAAATT 3450
```

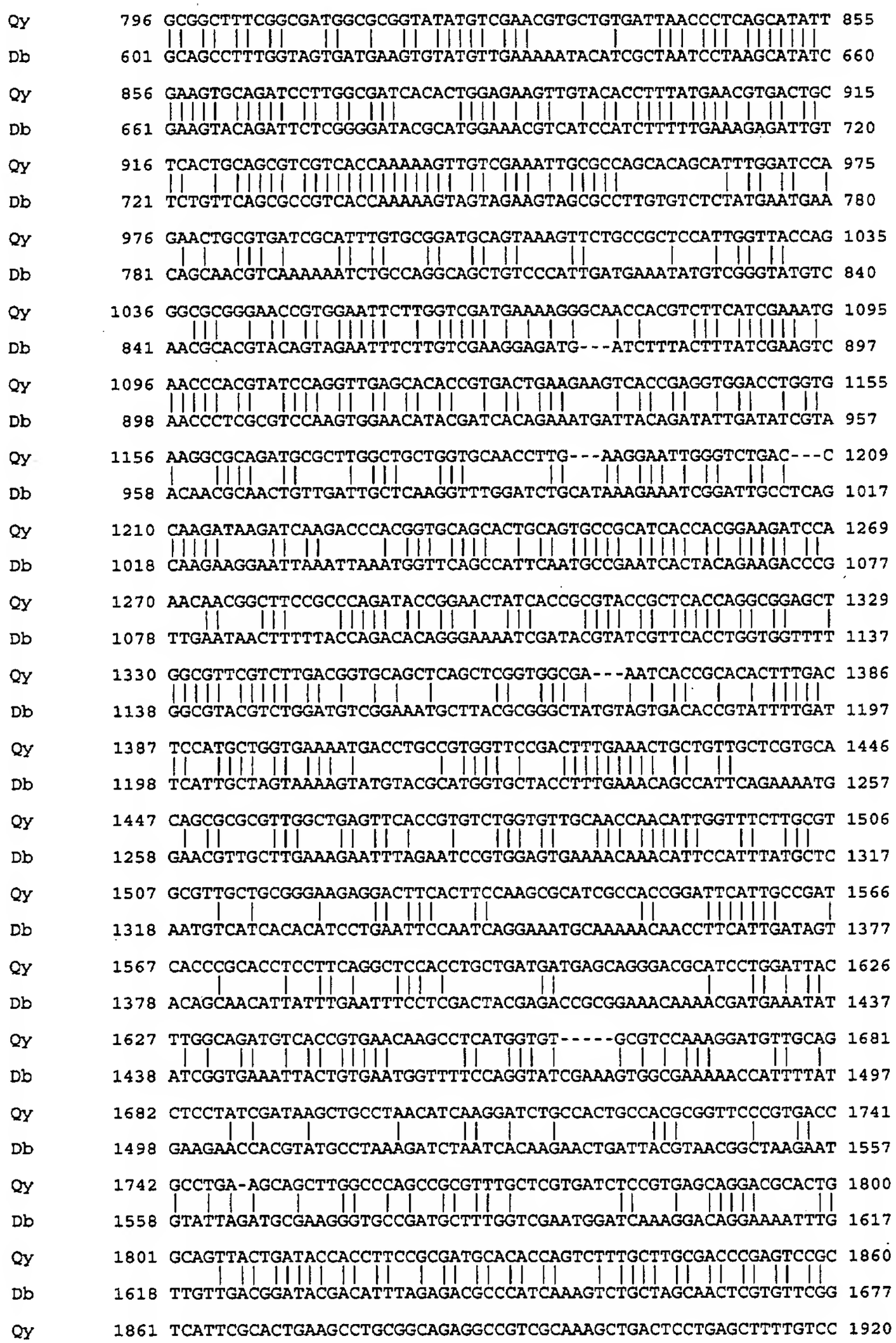
```
RESULT 8
US-08-956-171E-324/c
; Sequence 324, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 324:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5030 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 324:
US-08-956-171E-324

  Query Match           13.3%;  Score 479.8;  DB 4;  Length 5030;
  Best Local Similarity  48.3%;  Pred. No. 3.3e-126;
  Matches 1657;  Conservative    0;  Mismatches 1722;  Indels   54;  Gaps    9;

Qy        231 ATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGC 290
Db       3846 AATAAAAAAGTTACTTGTTGCTAACCGTGGAGAAATTGCAATTCGTATATTCAGAGCGGC 3787

Qy        291 ACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCA 350
Db       3786 GGCAGAATTAGACATCAGCACAGTTGCAATTTATTCGAATGAAGACAAAAGTTCATTACA 3727

Qy        351 CCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTA 410
Db       3726 TAGATATAAAGCAGATGAATCCTATTTAGTTGGAAGTGATTTAGGTCCTGCTGAAAGTTA 3667

Qy        411 CCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCC 470
Db       3666 TTTAAATATTGAGCGTATCATTGATGTAGCAAAACAAGCGAATGTGGATGCGATTCATCC 3607

Qy        471 GGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCAT 530
Db       3606 TGGCTATGGATTTTTAAGTGAAAATGAACAATTTGCGCGTCGTTGTGCAGAAGAAGGAAT 3547

Qy        531 TACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGT 590
Db       3546 TAAATTTATTGGTCCTCATTTAGAACATTTAGATATGTTTGGAGATAAAGTTAAAGCTCG 3487

Qy        591 AACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACAT 650
Db       3486 TACAACGGCTATCAAGGCAGATTTACCAGTTATTCCTGGTACAGACGGTCCAATTAAATC 3427

Qy        651 CGATGAGATCGTTAAAAGCGCTGAAGGCCAGAC---TTACCCCATCTTTGTGAAGGCAGT 707
Db       3426 ATATGAATTAGCAAAAGAATTTGCAGAAGAAGCTGGTTTCCCGCTAATGATTAAAGCCAC 3367

Qy        708 TGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATT 767
Db       3366 AAGTGGTGGCGGCGGTAAAGGTATGAGAATCGTTCGTGAAGAAAGTGAATTAGAAGATGC 3307

Qy        768 AGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGA 827
```

```
Db      3306 TTTCCATAGAGCAAAATCAGAAGCTGAAAAATCATTTGGTAATAGTGAAGTTTACATAGA 3247
Qy       828 ACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGA 887
Db      3246 AAGATACATTGATAATCCAAAGCATATTGAAGTACAAGTCATAGGTGACGAACATGGAAA 3187
Qy       888 AGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGA 947
Db      3186 TATCGTACACTTATTTGAACGTGATTGTTCAGTACAACGTCGTCATCAAAAAGTTGTAGA 3127
Qy       948 AATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGT 1007
Db      3126 AGTTGCACCATCAGTTGGATTATCATCAACATTACGTCAACGTATTTGTGATGCTGCAAT 3067
Qy      1008 AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGA 1067
Db      3066 TCAATTGATGGAAAATATTAAATATGTCAATGCAGGTACTGTTGAATTTCTAGTATCTG- 3008
Qy      1068 AAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGAC 1127
Db      3007 --GTGACGAATTCTTCTTTATAGAAGTTAACCCTCGTGTACAAGTAGAGCATACAATTAC 2950
Qy      1128 TGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAAC 1187
Db      2949 AGAGATGGTAACAGGAATTGATATTGTTAAGACACAAATTTTAGTTGCAGCAGGTGCCGA 2890
Qy      1188 CTTGAAGGAATTGG---------GTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGC 1238
Db      2889 TTTATTTGGTGAAGAGATTAATATGCCGCAACAAAAAGATATTACAACATTAGGCTATGC 2830
Qy      1239 ACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAAC 1298
Db      2829 CATCCAATGTCGTATTACAACAGAAGATCCGTTAAATGATTTCATGCCGGATACTGGAAC 2770
Qy      1299 TATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCT 1358
Db      2769 AATCATTGCTTATCGTTCAAGCGGGGGCTTTGGTGTACGTCTTGATGCTGGAGATGGTTT 2710
Qy      1359 C---GGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGG 1415
Db      2709 CCAAGGTGCTGAGATATCACCTTATTATGATTCATTACTCGTAAAATTATCTACACACGC 2650
Qy      1416 TTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTC 1475
Db      2649 GATATCATTTAAACAAGCAGAAGAAAAAATGGTACGCTCATTACGAGAAATGCGTATTCG 2590
Qy      1476 TGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTC 1535
Db      2589 TGGTGTTAAAACTAATATTCCATTCTTAATTAATGTAATGAAGAATAAAAAGTTCACAAG 2530
Qy      1536 CAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGC 1595
Db      2529 TGGTGATTACACAACTAAATTTATTGAAGAAACACCAGAACTTTTCGACATTCAGCCGTC 2470
Qy      1596 TGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAA--------C 1647
Db      2469 TCTAGATAGAGGTACTAAAACATTAGAATATATAGGTAATGTAACAATTAATGGTTTCCC 2410
Qy      1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
Db      2409 AAATGTTGAGAAACGTCCGAAACCAGACTATGAATTAGCATCAATTCCAACTGTATCTTC 2350
Qy      1708 AAG-GATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGC 1766
Db      2349 AAGTAAAATCGCTTCATTTAGTGGTACGAAACAATTGCTTGATGAAGTAGGTCCAAAAGG 2290
Qy      1767 GTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGA 1826
Db      2289 TGTAGCTGAATGGGTTAAAAAGCAGGATGATGTCTTACTAACAGATACAACCTTTAGAGA 2230
Qy      1827 TGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGA 1886
Db      2229 TGCACACCAATCATTATTAGCTACACGAGTTAGAACTAAGGATATGATTAATATCGCATC 2170
```

```
Qy      1887 GGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTA 1946
Db      2169 CAAAACAGCGGACGTATTTAAAGATGGTTTCTCACTAGAAATGTGGGGCGGTGCTACATT 2110
Qy      1947 CGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGA 2006
Db      2109 TGATGTGGCATATAATTTCTTGAAGGAAAACCCATGGGAACGACTTGAACGTCTACGTAA 2050
Qy      2007 GGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACAC 2066
Db      2049 AGCTATTCCAAATGTATTATTCCAAATGTTGTTACGTGCTTCAAACGCAGTTGGTTATAA 1990
Qy      2067 CCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGA 2126
Db      1989 AAACTATCCTGATAATGTTATTCATAAATTCGTACAAGAAAGTGCTAAAGCAGGCATAGA 1930
Qy      2127 CATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGC 2186
Db      1929 TGTCTTTAGAATTTTCGATTCATTAAACTGGGTAGATCAAATGAAAGTTGCCAATGAAGC 1870
Qy      2187 AGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTC 2246
Db      1869 AGTACAAGAAGCGGGCAAAATCTCAGAAGGTACTATTTGTTATACAGGTGACATTTTAAA 1810
Qy      2247 TGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGT 2306
Db      1809 TCCTGAGCGATCAAACATTTATACTTTAGAGTATTATGTCAAACTAGCTAAAGAGTTAGA 1750
Qy      2307 CAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGC 2366
Db      1749 ACGTGAAGGTTTCCATATTTTAGCGATTAAAGATATGGCAGGCTTATTAAAACCTAAAGC 1690
Qy      2367 GGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACAC 2426
Db      1689 CGCTTACGAATTGATTGGTGAGTTAAAATCAGCTGTAGATTTACCAATTCATCTTCACAC 1630
Qy      2427 CCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGA 2486
Db      1629 TCATGATACAAGTGGTAATGGTTTATTAACATACAAACAAGCAATAGATGCTGGTGTCGA 1570
Qy      2487 TGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGC 2546
Db      1569 TATCATTGATACTGCTGTTGCTTCAATGAGTGGTTTAACAAGTCAGCCAAGCGCCAATTC 1510
Qy      2547 CATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTC 2606
Db      1509 GTTATATTATGCATTAAATGGCTTCCCACGCCACCTTAGAACTGATATTGAAGGTATGGA 1450
Qy      2607 TGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAAC 2666
Db      1449 GTCACTTAGTCATTATTGGTCAACTGTACGTACTTATTATTCAGACTTTGAAAGTGATAT 1390
Qy      2667 CCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCT 2726
Db      1389 CAAATCACCGAATACTGAAATTTATCAACATGAAATGCCTGGTGGACAGTATTCGAATTT 1330
Qy      2727 GCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTA 2786
Db      1329 AAGTCAACAAGCTAAAAGTTTAGGTTTAGGCGAAAGATTTGATGAAGTCAAAGATATGTA 1270
Qy      2787 CGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGT 2846
Db      1269 TCGCAGAGTGAATTTCTTATTTGGTGATATCGTAAAAGTAACACCATCGTCTAAAGTAGT 1210
Qy      2847 TGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGA 2906
Db      1209 TGGTGATATGGCACTTTATATGGTACAAAATGATCTTGATGAACAATCCGTGATTACAGA 1150
Qy      2907 TCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAA 2966
Db      1149 TGGCTATAAATTAGATTTCCCAGAATCAGTAGTGTCGTTCTTCAAAGGTGAAATAGGACA 1090
```

```
Qy       2967 CCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGG 3026
Db       1089 ACCTGTAAATGGTTTTAATAAAGATTTACAAGCGGTTATTTTAAAAGGCCAAGAAGCACT 1030

Qy       3027 CAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTC 3086
Db       1029 AACAGCTCGTCCAGGTGAATATCTAGAGCCAGTTGATTTTGAAAAAGTCCGTGAGTTGCT 970

Qy       3087 CAAGGAACGTCGCAATAGCC--------------------TCAACCGCCTGCTGTTCCC 3125
Db        969 TGAAGAAGAGCAACAAGGTCCTGTTACGGAGCAAGATATTATTAGTTATGTACTATATCC 910

Qy       3126 GAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGA 3185
Db        909 AAAAGTATATGAACAATATATTCAAACTAGAAATCAATACGGAAACTTATCGTTACTTGA 850

Qy       3186 TGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGA 3245
Db        849 TACGCCTACATTCTTCTTTGGAATGCGTAATG---GTGAAACAGTAGAAATCGAAATCGA 793

Qy       3246 TGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTAT 3305
Db        792 TAAAGGTAAACGATTAATTATTAAACTAGAAACGATTAGTGAACCAGATGAAAATGGTAA 733

Qy       3306 GCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTC 3365
Db        732 TAGAACGATTTACTATGCGATGAATGGTCAAGCGAGACGTATTTACATTAAAGATGAAAA 673

Qy       3366 CGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGC 3425
Db        672 TGTGCATACAAATGCGAACGTTAAGCCAAAAGCAGATAAGAGTAATCCAAGTCATATCGG 613

Qy       3426 TGCACCATTCGCTGGTGTTGTCAC---CGTGACTGTTGCTGAAGGTGATGAGGTCAAGGC 3482
Db        612 TGCTCAAATGCCAGGTTCAGTAACTGAAGTCAAGGTTAGTGTAGGTGAAACTGTGAAAGC 553

Qy       3483 TGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGT 3542
Db        552 TAATCAGCCGTTGCTAATTACTGAAGCTATGAAAATGGAAACAACAATTCAAGCACCATT 493

Qy       3543 TGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTT 3602
Db        492 TGACGGTGTGATTAAACAAGTAACTGTAAATAATGGTGACACAATAGCGACAGGCGATTT 433

Qy       3603 GATCGTCGTCGTT 3615
Db        432 ATTAATCGAAATT 420
```

```
RESULT 9
US-09-107-532A-1156
; Sequence 1156, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 1156:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2922 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: circular
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...2922
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1156:
US-09-107-532A-1156

  Query Match             13.2%;  Score 476.2;  DB 4;  Length 2922;
  Best Local Similarity   49.9%;  Pred. No. 2.5e-125;
  Matches 1425;  Conservative    0;  Mismatches 1408;  Indels   24;  Gaps    8;

Qy        199 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 258
Db          1 GTTTATAATAAAAAAAGATGTAGGAGTGGTTCGATGAAAAAAGTTTTAGTTGCTAACAGG 60

Qy        259 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318
Db         61 GGTGAAATCGCTGTACGTGTATTTCGTGCTTGTACAGAATTAGGAATTAAAACTGTTGGA 120

Qy        319 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 378
Db        121 ATCTATGCAGAAAAAAATGAATATTCCGTACACCGTTTCAAAGCGGATGAAGCATATTTA 180

Qy        379 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 438
Db        181 GTAGGACAAGGAAAAAAACCGATTGATGCGTATCTAGATATTGAAGGGATCATTTCTATC 240

Qy        439 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 498
Db        241 GCGAAAGAATGCGGAGCTGACGCTATTCATCCGGGTTACGGTCTTTTGTCGGAAAACTTG 300

Qy        499 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 558
Db        301 AATTTTGCGCAACGGTGCGAAGAAAAAGGCATCATTTTTGTTGGACCCAAACTGCATCAT 360

Qy        559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGG---TCTG 615
Db        361 TTAGATATTTTTGGAGATAAAATCAAAGCAAAAGCTGCAGCTATCGAAGCGGGGATTGCT 420

Qy        616 CCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAA 675
Db        421 TCGATTCCAGGCACAGATGGACCAATCGCTTCTATAGACGATGCCTTGGAGTTTGCGAAA 480

Qy        676 GGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGT 735
Db        481 CAATACGGTTATCCGATTATGATCAAGGCAGCTCTATGCGGCGGCGGACGCGGCATGCGT 540

Qy        736 TTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAA 795
Db        541 GTAGCTCACGATGAAAAGAGTGCAAGAGAGGGCTACGAACGTGCGAAAAGCGAAGCAAAA 600
```


Qy 796 GCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATT 855
Db 601 GCAGCCTTTGGTAGTGATGAAGTGTATGTTGAAAAATACATCGCTAATCCTAAGCATATC 660
Qy 856 GAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGC 915
Db 661 GAAGTACAGATTCTCGGGGATACGCATGGAAACGTCATCCATCTTTTTGAAAGAGATTGT 720
Qy 916 TCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCA 975
Db 721 TCTGTTCAGCGCCGTCACCAAAAAGTAGTAGAAGTAGCGCCTTGTGTCTCTATGAATGAA 780
Qy 976 GAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAG 1035
Db 781 CAGCAACGTCAAAAAATCTGCCAGGCAGCTGTCCCATTGATGAAATATGTCGGGTATGTC 840
Qy 1036 GGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATG 1095
Db 841 AACGCACGTACAGTAGAATTTCTTGTCGAAGGAGATG---ATCTTTACTTTATCGAAGTC 897
Qy 1096 AACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTG 1155
Db 898 AACCCTCGCGTCCAAGTGGAACATACGATCACAGAAATGATTACAGATATTGATATCGTA 957
Qy 1156 AAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTG---AAGGAATTGGGTCTGAC---C 1209
Db 958 ACAACGCAACTGTTGATTGCTCAAGGTTTGGATCTGCATAAAGAAATCGGATTGCCTCAG 1017
Qy 1210 CAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCA 1269
Db 1018 CAAGAAGGAATTAAATTAAATGGTTCAGCCATTCAATGCCGAATCACTACAGAAGACCCG 1077
Qy 1270 AACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCT 1329
Db 1078 TTGAATAACTTTTTACCAGACACAGGGAAAATCGATACGTATCGTTCACCTGGTGGTTTT 1137
Qy 1330 GGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGA---AATCACCGCACACTTTGAC 1386
Db 1138 GGCGTACGTCTGGATGTCGGAAATGCTTACGCGGGCTATGTAGTGACACCGTATTTTGAT 1197
Qy 1387 TCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCA 1446
Db 1198 TCATTGCTAGTAAAAGTATGTACGCATGGTGCTACCTTTGAAACAGCCATTCAGAAAATG 1257
Qy 1447 CAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGT 1506
Db 1258 GAACGTTGCTTGAAAGAATTTAGAATCCGTGGAGTGAAAACAAACATTCCATTTATGCTC 1317
Qy 1507 GCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGAT 1566
Db 1318 AATGTCATCACACATCCTGAATTCCAATCAGGAAATGCAAAAACAACCTTCATTGATAGT 1377
Qy 1567 CACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTAC 1626
Db 1378 ACAGCAACATTATTTGAATTTCCTCGACTACGAGACCGCGGAAACAAAACGATGAAATAT 1437
Qy 1627 TTGGCAGATGTCACCGTGAACAAGCCTCATGGTGT-----GCGTCCAAAGGATGTTGCAG 1681
Db 1438 ATCGGTGAAATTACTGTGAATGGTTTTCCAGGTATCGAAAGTGGCGAAAAACCATTTTAT 1497
Qy 1682 CTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACC 1741
Db 1498 GAAGAACCACGTATGCCTAAAGATCTAATCACAAGAACTGATTACGTAACGGCTAAGAAT 1557
Qy 1742 GCCTGA-AGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db 1558 GTATTAGATGCGAAGGGTGCCGATGCTTTGGTCGAATGGATCAAAGGACAGGAAAATTTG 1617
Qy 1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db 1618 TTGTTGACGGATACGACATTTAGAGACGCCCATCAAAGTCTGCTAGCAACTCGTGTTCGG 1677
Qy 1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db 1678 ACAAAGGATTTCAAGCAGATTGCACGTTTGACTGGAGAAGGTCTTCCTGAACTCTTTTCA 1737
Qy 1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db 1738 AGTGAGATGTGGGGGGGAGCAACATTTGATGTGGCTTACCGCTTTTTGAATGAAGATCCT 1797
Qy 1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db 1798 TGGCAACGGTTGAGAAAAATTCGTTCTTTAATGCCTAATACGTTGCTGCAAATGCTTTTC 1857
Qy 2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db 1858 AGAGGTTCAAATGCCGTTGGGTATTCTAATTATCCTGACAATGTATTGGTGGAATTTGTC 1917
Qy 2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db 1918 AAAGAAGCAGCAGCGCAAGGAATCGATGTATTTCGAATCTTTGATAGTTTGAACTGGACT 1977
Qy 2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db 1978 CCACAAATGGAAAAAAGTATCCAAGCTGTACGGGATACTGGTA---AAATCGCAGAGGCA 2034
Qy 2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db 2035 GCTATTTGCTATACTGGCGATATCAACGATCCAAGTCGGGCAAAATACAATGTTCAGTAT 2094
Qy 2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db 2095 TATAAAGACATGGCAAAGGAACTAGAACAACTGGGTGCGCATATCATTGCAATCAAAGAT 2154
Qy 2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db 2155 ATGGCTGGGCTGCTTAAACCGCAGGCAGCTTATCGTTTGATTAGTGAATTGAAAGAAACG 2214
Qy 2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db 2215 ACAGATCTGCCAATCCATCTGCATACTCATGATACCAGCGGGAATGGAATCATCACTTAT 2274
Qy 2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db 2275 TCTGCTGCTTCAAAAGCAGGAGTAGACATCGTTGACGTGGCAATGAGTGCAATGAGTGGG 2334
Qy 2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db 2335 AATACTAGTCAGCCTAGCATGAGCAGCTTGTATTATGCACTGGTCAATGGTCCACGGCTG 2394
Qy 2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db 2395 CCAGAAATCACGATTGAAAATGCACAGAAATTAAACCATTATTGGGAAGATGTGCGCATG 2454
Qy 2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db 2455 TATTACAAACCGTTTGAAAACGGCTTGAATGCTCCTGAAACAGAAGTATATATGCATGAA 2514
Qy 2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db 2515 ATGCCAGGAGGACAGTATTCGAATCTACAGCAACAGGCAAAAGCTGTCGGATTAGGTCAT 2574
Qy 2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db 2575 CGCTGGGATGAAATCAAACAAATGTATCATACTGTCAACTTGATGTTTGGCGATATCGTT 2634
Qy 2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db 2635 AAAGTAACTCCATCATCTAAAGTAGTTGGAGATATGGCACTATTTATGGTCCAAAATGAT 2694
Qy 2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db 2695 CTTACAGAAGAAGATATTTATGAAAAAGGCGAAACACTAAGTTTCCCTGAATCTGTTGTG 2754
Qy 2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000

```
Db       2755 ACTTTCTTCCAAGGAGAACTAGGGCAACCCGTTGGCGGATTTCCTAAAAAATTGCAAAAG 2814

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTC 3037
                ||  ||||| |  | | |   || | || ||
Db       2815 ATTATTCTTAAAGGCAGACCTGCATTGAATGAACGTC 2851


RESULT 10
US-09-543-681A-1699
; Sequence 1699, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1699
;   LENGTH: 1362
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-09-543-681A-1699

  Query Match             8.7%;  Score 313.4;  DB 4;  Length 1362;
  Best Local Similarity   55.4%;  Pred. No. 4.4e-79;
  Matches  721;  Conservative    0;  Mismatches  556;  Indels   24;  Gaps    5;

Qy        236 AAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCG 295
              |||| ||| |  | || ||||| || ||||| ||  | |||    |  | ||       |
Db         20 AAAAAATCCTCATTGCCAACCGTGGTGAAATTGCACTGCGTATCCTAAGAGCTTGTAAGG 79

Qy        296 AAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCT 355
              ||   ||         | || ||  || || ||   | ||| || |  | |   |||
Db         80 AACTTGGGATCAAAGCAGTCGCCGTTCACTCCACGGCAGACCGTGATTTAAAACACGTTC 139

Qy        356 CTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGG 415
              | ||    ||  || || | |||||| |||  | ||| ||   ||    ||| |
Db        140 TGCTGGCAGACGAGACTATCTGTATTGGTCCCGCTGCTTCAGCA---AAAAGTTACTTAA 196

Qy        416 ACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGAT 475
              | ||    | |||||| |  || ||  | | |   ||  | |||||| |||| ||||
Db        197 ATATTCCGGCAATTATTGCCGCGGCAGAGATAAGTGGCGCGCAAGCCATTCACCCAGGAT 256

Qy        476 ACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTT 535
              | |||||||||||||||||||| |  |||||    |  ||          ||| ||| ||
Db        257 ATGGCTTCCTGTCTGAAAATGCCGATTTTGCCGAACAAGTTGAACGCTCAGGCTTTATTT 316

Qy        536 TTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCG 595
              ||||||||||  |   | ||    ||  || |  ||||||||    |  |||  |    |
Db        317 TTATTGGCCCTAAAGCGGAAACCATTCGCCTAATGGGTGATAAAGTTTCCGCTATTGAAG 376

Qy        596 CCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATC------CACCCCGAGCAAAAACA 649
              |   ||| || || ||| | ||   | |  | |  ||      | |    | ||  | |
Db        377 CGATGAAAAAAGCCGGTGTTCCTTGTGTACCAGGCTCAGATGGCCCATTAGGTAACGATA 436

Qy        650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
              |   | |   | ||| |||  || ||      ||||||  |  |  | || |||   |
Db        437 CAGCCAAAAATATCGAAATCGCCAAACGCATTGGTTACCCTGTTATCATCAAAGCATCAG 496

Qy        710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
              |||||  ||  || |||||||||||| |||||   |||    | ||  |     ||  |
Db        497 GTGGTGGCGGTGGTCGCGGTATGCGTGTTGTTCGCTCAGAAAAAGACTTAGCGCAAGCAA 556

Qy        770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
              |   |   |   |||||  |||||||  ||  || |  |   ||| ||| | |||
Db        557 TCTCCATGACCCGTGCGGAAGCCAAAGCGGCATTTAGCAACGATATGGTCTATATGGAAA 616

Qy        830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
                  |    || || |  ||  | ||| | |||  |  | | ||||      ||  | |
```

```
Db        617 AGTACCTTGAAAATCCACGCCACGTCGAAATTCAGGTGATGGCCGATGGACAAGGTAATG 676

Qy        890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
              | | |  |  |||||||||||||||| |||| ||||| |||||||||| || |||
Db        677 CTATCTATTTAGCTGAACGTGACTGCTCAATGCAACGTCGCCACCAAAAAGTGGTTGAAG 736

Qy        950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
              || |||||||  |  | |    || ||  | ||| |    ||  |||        | ||
Db        737 AAGCACCAGCACCGGGTATTACCCCTGAGATCCGTAAAAATATCGGTGAACGCTGTGCAA 796

Qy       1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
              |  |||      ||||| |||| || |||||  ||  | ||||||    || ||||||
Db        797 ATGCCTGTATTGAAATTGGCTACCGCGGTGCGGGTACGTTTGAATTC---CTCTATGAAA 853

Qy       1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
              | |||  |   ||| |||||||||||  | |||||  ||||||||||  | || ||||
Db        854 ATGGCGAATTCTACTTTATCGAAATGAATACCCGTATTCAGGTTGAGCATCCTGTTACTG 913

Qy       1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
              |      ||||||  || |||||  | || | |||  |||  | ||  | ||  | ||
Db        914 AGATGATCACCGGTGTTGACCTTATCAAAGAGCAACTGCGTATTGCATCAGGCTTACC-- 971

Qy       1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
                    |||     | || |||||| | ||  |     || ||   ||  |  ||||||
Db        972 -------ATTATCAGTCACGCAAGATCAAATTCACGTTCATGGACATGCTATTGAGTGCC 1024

Qy       1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
              | |||| | | |||||||||||| |    ||||   |||    | |||||  ||||  |
Db       1025 GTATCAACGCAGAAGATCCAAAAA---CCTTCTTGCCAAGCCCGGGAACCATCACTCGTT 1081

Qy       1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAA 1369
              || |||||||||||||  |||  || |||   ||          | |   |||  |
Db       1082 TCCACTCACCAGGCGGATTTGGTGTACGTTGGGAATCACATATTTACGCAGGTTACACCG 1141

Qy       1370 TCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAA 1429
              |  | ||||||  ||| ||  |||  |  |||  |||  |      |||          |||||
Db       1142 TTCCACCACACTATGATTCAATGATTGGTAAATTGATCACTTACGGTGAAACACGTGAAA 1201

Qy       1430 CTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCA 1489
              |||  || |||||    |    |||||||| ||  | |    |    ||  ||  |||||
Db       1202 TTGCGATTTCTCGTATGAAAAATGCGTTGGCAGAACTTATTATTGACGGCATTAAAACCA 1261

Qy       1490 ACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTC 1530
              | ||||  |    |      |  ||   || ||| | |||
Db       1262 ATATTGAGCTACACCAATTCATTATGAATGATGAGCAATTC 1302


RESULT 11
US-07-956-700B-5
; Sequence 5, Application US/07956700B
; Patent No. 5539092
;  GENERAL INFORMATION:
;    APPLICANT:  Robert Haselkorn and Piotr Gornicki
;    TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;    TITLE OF INVENTION:  Carboxylase
;    NUMBER OF SEQUENCES:  116
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  321 No. 5539092th Clark Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60610
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/956,700B
;      FILING DATE:  19921002
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Thomas E. No. 5539092thrup
;      REGISTRATION NUMBER:  33,268
;      REFERENCE/DOCKET NUMBER:  ARCD:058
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  1-312-744-0090
;      TELEFAX:  1-312-755-4489
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1362 base pairs
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Oligonucleotide
US-07-956-700B-5

  Query Match             8.6%;  Score 313;  DB 1;  Length 1362;
  Best Local Similarity   53.9%;  Pred. No. 5.7e-79;
  Matches  718;  Conservative    0;  Mismatches  600;  Indels   15;  Gaps    3;

Qy       232 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCA 291
Db         7 TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy       292 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCAC 351
Db        67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy       352 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTAC 411
Db       127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy       412 CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCG 471
Db       184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy       472 GGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATT 531
Db       244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy       532 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC---G 588
Db       304 ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG 363

Qy       589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
Db       364 GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy       649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
Db       424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy       709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
Db       484 GCGGGGGGCGGTGGTCGCGGTATGCGGCTGGTGCGTGACCCTGCAGATCTGGAAAAACTG 543

Qy       769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
Db       544 TTCCTTGCTGCCCAAGGAGAAGCCGAGGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy       829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
Db       604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy       889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
Db       664 GTAGTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy       949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
Db       724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCGCCGTC 783

Qy      1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
Db       784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy      1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
Db       844 ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy      1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
Db       904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy      1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
Db       959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy      1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
Db      1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCATTACAGGC 1074

Qy      1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
Db      1075 TATTTACCGCCCGGCGGCCCCGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134

Qy      1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
Db      1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy      1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
Db      1195 GAGGCGATCGCGCGGATGCAGCGTGCTCTGCGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy      1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
Db      1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy      1549 ACCGGATTCATTG 1561
Db      1315 ACCAACTTTGTTG 1327


RESULT 12
US-08-476-537-5
; Sequence 5, Application US/08476537
; Patent No. 5756290
;  GENERAL INFORMATION:
;    APPLICANT:  Robert Haselkorn and Piotr Gornicki
;    TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;    TITLE OF INVENTION:  Carboxylase
;    NUMBER OF SEQUENCES:  116
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  321 No. 5756290th Clark Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60610
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,537
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/956,700
;      FILING DATE:  10/21/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas E. No. 5756290thrup
;      REGISTRATION NUMBER:  33,268
;      REFERENCE/DOCKET NUMBER:  ARCD:058
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  1-312-744-0090
```

```
;      TELEFAX:  1-312-755-4489
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1362 base pairs
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Oligonucleotide
US-08-476-537-5

  Query Match             8.6%;  Score 313;  DB 1;  Length 1362;
  Best Local Similarity   53.9%;  Pred. No. 5.7e-79;
  Matches  718;  Conservative    0;  Mismatches  600;  Indels   15;  Gaps    3;

Qy        232 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCA 291
              ||||| |||||| || | || || ||||||||||||||  | ||   | ||||  |
Db          7 TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy        292 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCAC 351
              |  |||  |||    | |||| | ||  || || ||  ||  |||||     | ||||
Db         67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy        352 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTAC 411
                || ||    ||||| ||| | |||||   |||||   |     |||    ||
Db        127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy        412 CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCG 471
              || | |||    | || || | || ||     |   ||| ||   ||||| ||||
Db        184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy        472 GGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATT 531
              || || |||||| || | || |||||||   ||||     || || ||  ||     |
Db        244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy        532 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC---G 588
              || ||||||||||| | ||| ||     ||     ||  || ||||| ||   |||  |
Db        304 ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG 363

Qy        589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
              | |||      |  |   | | || | |  ||  ||       |  | |   |   |
Db        364 GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy        649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
              | |||  |   | |||   ||||  |         || ||| || |  | ||  ||
Db        424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy        709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
              || || || || || ||||||||||||  | ||   |  ||||  || ||    ||| |
Db        484 GCGGGGGGCGGTGGTCGCGGTATGCGGCTGGTGCGTGACCCTGCAGATCTGGAAAAACTG 543

Qy        769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
                |  ||    | ||||| || || ||||| ||  ||   |   | ||| |||||
Db        544 TTCCTTGCTGCCCAAGGAGAAGCCGAGGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy        829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
                |  |   | ||| |  || ||||| | |||||| | |  |||| |    ||  |
Db        604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy        889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
              || || ||  ||    || || || ||||| | || ||||||||||||||  |  |||||
Db        664 GTAGTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy        949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
                || ||   | |    |  ||||  ||||| |    ||  | |  |  || ||
Db        724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCGCCGTC 783

Qy       1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
              ||  ||   |  | || || |||   || || || |||||||| ||  |||||||||
Db        784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy       1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
              |  ||||||  |  |||||| || |||||  | || |||||  || |||||  | || ||
Db        844 ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy       1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
              |||    | || |  |||| || |  || |||||| ||  | ||    ||| |
Db         904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy       1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
                | |  || |  |   ||||  | ||  |   | |||    ||  |  | |||
Db         959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy       1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
              || ||||   |||||||||||  |  ||   ||||| ||  || | ||    || ||| |
Db        1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCATTACAGGC 1074

Qy       1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
              ||      | || |||||  | |||||||||| | ||      | | |  ||   ||||
Db        1075 TATTTACCGCCCGGCGGCCCCGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134

Qy       1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
              ||  |  |  |  || ||  || | |  ||| |||     | | ||| |      |||
Db        1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy       1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
                ||  | ||  ||    |||||  ||  ||   ||  |  ||  |  |||  |  | ||
Db        1195 GAGGCGATCGCGCGGATGCAGCGTGCTCTGCGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy       1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
              | |  || |||||    |   || |||  |      ||  |||     |  |    ||
Db        1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy       1549 ACCGGATTCATTG 1561
              |||   ||  |||
Db        1315 ACCAACTTTGTTG 1327


RESULT 13
US-08-485-607-5
; Sequence 5, Application US/08485607
; Patent No. 5792627
;  GENERAL INFORMATION:
;    APPLICANT:  Robert Haselkorn and Piotr Gornicki
;    TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;    TITLE OF INVENTION:  Carboxylase
;    NUMBER OF SEQUENCES:  116
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  321 No. 5792627th Clark Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60610
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,607
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/956,700
;      FILING DATE:  10/21/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas E. No. 5792627thrup
;      REGISTRATION NUMBER:  33,268
;      REFERENCE/DOCKET NUMBER:  ARCD:058
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  1-312-744-0090
;      TELEFAX:  1-312-755-4489
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1362 base pairs
;      TYPE:  Nucleic acid
```

```
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Oligonucleotide
US-08-485-607-5

  Query Match              8.6%;  Score 313;  DB 1;  Length 1362;
  Best Local Similarity   53.9%;  Pred. No. 5.7e-79;
  Matches  718;  Conservative    0;  Mismatches  600;  Indels   15;  Gaps    3;

Qy        232 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCA 291
Db          7 TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy        292 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCAC 351
Db         67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy        352 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTAC 411
Db        127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy        412 CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCG 471
Db        184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy        472 GGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATT 531
Db        244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy        532 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC---G 588
Db        304 ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG 363

Qy        589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
Db        364 GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy        649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
Db        424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy        709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
Db        484 GCGGGGGGCGGTGGTCGCGGTATGCGGCTGGTGCGTGACCCTGCAGATCTGGAAAAACTG 543

Qy        769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
Db        544 TTCCTTGCTGCCCAAGGAGAAGCCGAGGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy        829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
Db        604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy        889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
Db        664 GTAGTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy        949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
Db        724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCGCCGTC 783

Qy       1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
Db        784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy       1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
Db        844 ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy       1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
Db        904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy       1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
Db        959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy       1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
Db       1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCATTACAGGC 1074

Qy       1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
Db       1075 TATTTACCGCCCGGCGGCCCCGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134

Qy       1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
Db       1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy       1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
Db       1195 GAGGCGATCGCGCGGATGCAGCGTGCTCTGCGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy       1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
Db       1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy       1549 ACCGGATTCATTG 1561
Db       1315 ACCAACTTTGTTG 1327

RESULT 14
US-08-475-879-5
; Sequence 5, Application US/08475879
; Patent No. 5972644
; Patent No. 5972644 5786170
;  GENERAL INFORMATION:
;    APPLICANT:  Robert Haselkorn and Piotr Gornicki
;    TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;    TITLE OF INVENTION:  Carboxylase
;    NUMBER OF SEQUENCES:  116
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  321 No. 5972644 5786170th Clark Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60610
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,879
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/956,700
;      FILING DATE:  10/21/92
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas E. No. 5972644 5786170thrup
;      REGISTRATION NUMBER:  33,268
;      REFERENCE/DOCKET NUMBER:  ARCD:058
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  1-312-744-0090
;      TELEFAX:  1-312-755-4489
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1362 base pairs
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Oligonucleotide
US-08-475-879-5
```

```
Query Match             8.6%;  Score 313;  DB 2;  Length 1362;
Best Local Similarity   53.9%;  Pred. No. 5.7e-79;
Matches  718;  Conservative    0;  Mismatches  600;  Indels   15;  Gaps    3;

Qy        232 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCA 291
Db          7 TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy        292 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCAC 351
Db         67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy        352 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTAC 411
Db        127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy        412 CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCG 471
Db        184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy        472 GGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATT 531
Db        244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy        532 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC---G 588
Db        304 ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG 363

Qy        589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
Db        364 GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy        649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
Db        424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy        709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
Db        484 GCGGGGGGCGGTGGTCGCGGTATGCGGCTGGTGCGTGACCCTGCAGATCTGGAAAAACTG 543

Qy        769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
Db        544 TTCCTTGCTGCCCAAGGAGAAGCCGAGGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy        829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
Db        604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy        889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
Db        664 GTAGTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy        949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
Db        724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCGCCGTC 783

Qy       1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
Db        784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy       1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
Db        844 ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy       1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
Db        904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy       1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
Db        959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy       1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
Db       1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCATTACAGGC 1074

Qy       1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
Db       1075 TATTTACCGCCCGGCGGCCCCGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134

Qy       1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
Db       1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy       1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
Db       1195 GAGGCGATCGCGCGGATGCAGCGTGCTCTGCGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy       1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
Db       1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy       1549 ACCGGATTCATTG 1561
Db       1315 ACCAACTTTGTTG 1327
```

```
RESULT 15
US-09-433-043B-5
; Sequence 5, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
;  APPLICANT: HASELKORN, ROBERT
;  APPLICANT:  GORNICKI, PIOTR
;  TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
;  FILE REFERENCE: ARCD:338US
;  CURRENT APPLICATION NUMBER: US/09/433,043B
;  CURRENT FILING DATE:  1999-10-25
;  PRIOR APPLICATION NUMBER: 08/475,879
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: 07/956,700
;  PRIOR FILING DATE: 1992-10-02
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1362
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-433-043B-5

  Query Match             8.6%;  Score 313;  DB 4;  Length 1362;
  Best Local Similarity   53.9%;  Pred. No. 5.7e-79;
  Matches  718;  Conservative    0;  Mismatches  600;  Indels   15;  Gaps    3;

Qy        232 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCA 291
Db          7 TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy        292 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCAC 351
Db         67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy        352 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTAC 411
Db        127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy        412 CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCG 471
Db        184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy        472 GGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATT 531
Db        244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303
```

```
Qy       532 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGC---G 588
Db       304 ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG 363

Qy       589 GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAC 648
Db       364 GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy       649 ATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTT 708
Db       424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy       709 GCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTA 768
Db       484 GCGGGGGGCGGTGGTCGCGGTATGCGGCTGGTGCGTGACCCTGCAGATCTGGAAAAACTG 543

Qy       769 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAA 828
Db       544 TTCCTTGCTGCCCAAGGAGAAGCCGAGGCAGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy       829 CGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAA 888
Db       604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy       889 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 948
Db       664 GTAGTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy       949 ATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTA 1008
Db       724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCGCCGTC 783

Qy      1009 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 1068
Db       784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy      1069 AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 1128
Db       844 ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy      1129 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC 1188
Db       904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy      1189 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1248
Db       959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy      1249 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1308
Db      1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCATTACAGGC 1074

Qy      1309 TACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1368
Db      1075 TATTTACCGCCCGGCGGCCCCGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134

Qy      1369 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1428
Db      1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy      1429 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1488
Db      1195 GAGGCGATCGCGCGGATGCAGCGTGCTCTGCGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy      1489 AACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1548
Db      1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy      1549 ACCGGATTCATTG 1561
Db      1315 ACCAACTTTGTTG 1327
```

Search completed: March 23, 2004, 18:39:49
Job time : 257 secs

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 20:24:42 ; Search time 1213 Seconds
(without alignments)
11042.568 Million cell updates/sec

Title: US-10-045-072-1
Perfect score: 3621
Sequence: 1 tggggcggggttagatcctg.....tgatcgtcgtcgtttcctaa 3621

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3621	100.0	3621	14	US-10-045-072-1	Sequence 1, Appli
2	3621	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appli
3	3420	94.4	3420	9	US-09-738-626-765	Sequence 765, App
4	3398.8	93.9	3474	9	US-09-974-973-1	Sequence 1, Appli
5	3398.8	93.9	3474	9	US-09-974-973-3	Sequence 3, Appli
6	1927.4	53.2	3423	12	US-10-282-122A-17701	Sequence 17701, A
7	1361.6	37.6	3381	12	US-10-282-122A-25766	Sequence 25766, A
8	1294.4	35.7	3384	12	US-10-282-122A-28626	Sequence 28626, A
9	1292.2	35.7	3381	12	US-10-282-122A-26500	Sequence 26500, A
10	681	18.8	3945	9	US-09-917-800A-1566	Sequence 1566, Ap
11	677.4	18.7	4017	9	US-09-880-107-3029	Sequence 3029, Ap
12	659	18.2	3432	15	US-10-369-493-31453	Sequence 31453, A
13	643.6	17.8	3459	15	US-10-369-493-35714	Sequence 35714, A
14	617	17.0	3441	15	US-10-369-493-35137	Sequence 35137, A
15	617	17.0	3453	15	US-10-369-493-38504	Sequence 38504, A
16	617	17.0	3453	15	US-10-369-493-38693	Sequence 38693, A
17	601.8	16.6	3453	15	US-10-369-493-41025	Sequence 41025, A
18	598	16.5	4152	15	US-10-398-221-3692	Sequence 3692, Ap
19	582	16.1	3414	15	US-10-369-493-42038	Sequence 42038, A
20	577.2	15.9	3441	12	US-10-282-122A-24304	Sequence 24304, A
21	572.8	15.8	3447	15	US-10-369-493-46799	Sequence 46799, A
22	565.2	15.6	3537	15	US-10-369-493-45626	Sequence 45626, A
23	565	15.6	1163020	15	US-10-398-221-10	Sequence 10, Appl
24	565	15.6	3011208	15	US-10-398-221-2058	Sequence 2058, Ap
25	549.4	15.2	3540	15	US-10-369-493-25178	Sequence 25178, A
26	548.8	15.2	3501	15	US-10-369-493-46506	Sequence 46506, A
27	548.8	15.2	3555	15	US-10-369-493-26175	Sequence 26175, A
28	547.2	15.1	2709	15	US-10-369-493-37987	Sequence 37987, A
29	522	14.4	3444	12	US-10-282-122A-10682	Sequence 10682, A
30	514	14.2	3429	9	US-09-815-242-6709	Sequence 6709, Ap
31	509.2	14.1	3426	12	US-10-282-122A-6344	Sequence 6344, Ap
c 32	500.4	13.8	20072	9	US-09-070-927A-89	Sequence 89, Appl
33	482.4	13.3	3426	12	US-10-282-122A-21758	Sequence 21758, A
c 34	479.8	13.3	5030	8	US-08-781-986A-324	Sequence 324, App
c 35	479.8	13.3	5030	12	US-10-329-624-324	Sequence 324, App
36	479.6	13.2	3453	12	US-10-282-122A-8207	Sequence 8207, Ap
37	478.2	13.2	3441	9	US-09-815-242-4413	Sequence 4413, Ap
38	462.6	12.8	3435	12	US-10-282-122A-15904	Sequence 15904, A
39	449.4	12.4	3432	12	US-10-282-122A-16593	Sequence 16593, A
40	421.8	11.6	3222	9	US-09-815-242-8263	Sequence 8263, Ap
41	420.6	11.6	3429	12	US-10-282-122A-16733	Sequence 16733, A
42	363.8	10.0	1362	15	US-10-369-493-36893	Sequence 36893, A
43	358.4	9.9	3230	15	US-10-369-493-37518	Sequence 37518, A
44	337.2	9.3	4161	15	US-10-369-493-27492	Sequence 27492, A
45	311.4	8.6	1362	9	US-09-767-479-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-045-072-1
; Sequence 1, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-10-045-072-1

Query Match 100.0%; Score 3621; DB 14; Length 3621;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60

Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
```

```
Qy       121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
Db       121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
Qy       181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
Db       181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
Qy       241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
Db       241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
Qy       301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Db       301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Qy       361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Db       361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Qy       421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Db       421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Qy       481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Db       481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Qy       541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Db       541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Qy       601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Db       601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Qy       661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Db       661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Qy       721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Db       721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Qy       781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Db       781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Qy       841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Db       841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Qy       901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Db       901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Qy       961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Db       961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Qy      1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Db      1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Qy      1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Db      1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Qy      1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Db      1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
```

```
Qy      1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Db      1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Qy      1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Db      1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Qy      1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Db      1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Qy      1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Db      1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Qy      1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Db      1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Qy      1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Db      1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Qy      1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Db      1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Qy      1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Db      1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Qy      1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Db      1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Qy      1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db      1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Qy      1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db      1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Qy      1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db      1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Qy      1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db      1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Qy      1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db      1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Qy      2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db      2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Qy      2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db      2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Qy      2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db      2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Qy      2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db      2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Qy      2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
```

```
Db       2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Qy       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db       2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Qy       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db       2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Db       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Db       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Db       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Db       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Db       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Db       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Db       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Db       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Db       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Db       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Db       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Db       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Db       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Db       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Db       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Db       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Db       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Db       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Db       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
Db       3601 TTGATCGTCGTCGTTTCCTAA 3621
```

RESULT 2
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 100.0%; Score 3621; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 60
Db     705013 TGGGGCGGGGTTAGATCCTGGGGGGTTTATTTCATTCACTTTGGCTTGAAGTCGTGCAGG 705072
Qy         61 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 120
Db     705073 TCAGGGGAGTGTTGCCCGAAAACATTGAGAGGAAAACAAAAACCGATGTTTGATTGGGGG 705132
Qy        121 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 180
Db     705133 AATCGGGGGTTACGATACTAGGACGCAGTGACTGCTATCACCCTTGGCGGTCTCTTGTTG 705192
Qy        181 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 240
Db     705193 AAAGGAATAATTACTCTAGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAG 705252
Qy        241 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 300
```

```
Db    705253 ATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACC 705312
Qy       301 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 360
Db    705313 GGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTT 705372
Qy       361 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 420
Db    705373 GCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 705432
Qy       421 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 480
Db    705433 GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGC 705492
Qy       481 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 540
Db    705493 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATT 705552
Qy       541 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 600
Db    705553 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 705612
Qy       601 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 660
Db    705613 AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC 705672
Qy       661 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 720
Db    705673 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 705732
Qy       721 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 780
Db    705733 GGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 705792
Qy       781 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 840
Db    705793 TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT 705852
Qy       841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
Db    705853 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 705912
Qy       901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
Db    705913 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 705972
Qy       961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
Db    705973 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 706032
Qy      1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
Db    706033 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 706092
Qy      1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
Db    706093 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 706152
Qy      1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
Db    706153 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 706212
Qy      1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
Db    706213 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 706272
Qy      1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
Db    706273 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 706332
Qy      1321 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 1380
Db    706333 GGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC 706392
Qy      1381 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 1440
Db    706393 TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCT 706452
Qy      1441 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 1500
Db    706453 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC 706512
Qy      1501 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 1560
Db    706513 TTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT 706572
Qy      1561 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 1620
Db    706573 GCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTG 706632
Qy      1621 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 1680
Db    706633 GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA 706692
Qy      1681 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 1740
Db    706693 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGAC 706752
Qy      1741 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 1800
Db    706753 CGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTG 706812
Qy      1801 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 1860
Db    706813 GCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGC 706872
Qy      1861 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 1920
Db    706873 TCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCC 706932
Qy      1921 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 1980
Db    706933 GTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCG 706992
Qy      1981 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 2040
Db    706993 TGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTT 707052
Qy      2041 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 2100
Db    707053 CGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTT 707112
Qy      2101 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 2160
Db    707113 AAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTC 707172
Qy      2161 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 2220
Db    707173 TCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTG 707232
Qy      2221 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 2280
Db    707233 GCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTAC 707292
Qy      2281 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 2340
Db    707293 TACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGAT 707352
Qy      2341 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 2400
Db    707353 ATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAA 707412
Qy      2401 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 2460
Db    707413 TTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTAC 707472
```

```
Qy       2461 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707473 TTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGC 707532

Qy       2521 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707533 ACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 707592

Qy       2581 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707593 ACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGA 707652

Qy       2641 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707653 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAA 707712

Qy       2701 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707713 ATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGAT 707772

Qy       2761 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707773 CGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACC 707832

Qy       2821 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707833 AAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGT 707892

Qy       2881 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707893 GTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATC 707952

Qy       2941 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     707953 GCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACC 708012

Qy       3001 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708013 CGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAG 708072

Qy       3061 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708073 CAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTG 708132

Qy       3121 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708133 TTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCG 708192

Qy       3181 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708193 CTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTG 708252

Qy       3241 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708253 CCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAG 708312

Qy       3301 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708313 GGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGAC 708372

Qy       3361 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708373 CGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCAT 708432

Qy       3421 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708433 GTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAG 708492

Qy       3481 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708493 GCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCT 708552

Qy       3541 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     708553 GTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGAC 708612

Qy       3601 TTGATCGTCGTCGTTTCCTAA 3621
             |||||||||||||||||||||
Db     708613 TTGATCGTCGTCGTTTCCTAA 708633
```

RESULT 3
US-09-738-626-765
; Sequence 765, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 765
; LENGTH: 3420
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-765

Query Match 94.4%; Score 3420; DB 9; Length 3420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        199 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC 60

Qy        259 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCT 120

Qy        319 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC 180

Qy        379 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA 240

Qy        439 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC 300

Qy        499 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT 360

Qy        559 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCA 420
```

```
Qy       619 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 678
Db       421 GTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC 480

Qy       679 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 738
Db       481 CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTT 540

Qy       739 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 798
Db       541 GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG 600

Qy       799 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 858
Db       601 GCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAA 660

Qy       859 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 918
Db       661 GTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA 720

Qy       919 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 978
Db       721 CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAA 780

Qy       979 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 1038
Db       781 CTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGC 840

Qy      1039 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 1098
Db       841 GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC 900

Qy      1099 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 1158
Db       901 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG 960

Qy      1159 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1218
Db       961 GCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAG 1020

Qy      1219 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1278
Db      1021 ATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGC 1080

Qy      1279 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1338
Db      1081 TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT 1140

Qy      1339 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1398
Db      1141 CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG 1200

Qy      1399 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1458
Db      1201 AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTG 1260

Qy      1459 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1518
Db      1261 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1320

Qy      1519 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1578
Db      1321 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1380

Qy      1579 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1638
Db      1381 CTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTC 1440

Qy      1639 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1698
Db      1441 ACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1500

Qy      1699 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1758
Db      1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGC 1560

Qy      1759 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1818
Db      1561 CCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620

Qy      1819 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1878
Db      1621 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680

Qy      1879 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1938
Db      1681 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGC 1740

Qy      1939 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1998
Db      1741 GCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800

Qy      1999 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 2058
Db      1801 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTG 1860

Qy      2059 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 2118
Db      1861 GGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCC 1920

Qy      2119 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2178
Db      1921 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980

Qy      2179 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2238
Db      1981 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGT 2040

Qy      2239 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2298
Db      2041 GATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAG 2100

Qy      2299 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2358
Db      2101 GAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGC 2160

Qy      2359 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2418
Db      2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCAC 2220

Qy      2419 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2478
Db      2221 GTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCT 2280

Qy      2479 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2538
Db      2281 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2340

Qy      2539 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2598
Db      2341 CTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAG 2400

Qy      2599 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2658
Db      2401 GCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAG 2460

Qy      2659 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2718
Db      2461 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTG 2520

Qy      2719 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2778
Db      2521 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAA 2580

Qy      2779 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2838
```

```
Db       2581 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCC 2640
Qy       2839 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2898
Db       2641 AAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTT 2700
Qy       2899 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2958
Db       2701 GCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAG 2760
Qy       2959 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 3018
Db       2761 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGC 2820
Qy       3019 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 3078
Db       2821 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 2880
Qy       3079 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 3138
Db       2881 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAA 2940
Qy       3139 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3198
Db       2941 GAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTC 3000
Qy       3199 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3258
Db       3001 TTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3060
Qy       3259 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3318
Db       3061 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTG 3120
Qy       3319 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3378
Db       3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTC 3180
Qy       3379 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3438
Db       3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCT 3240
Qy       3439 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3498
Db       3241 GGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3300
Qy       3499 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3558
Db       3301 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360
Qy       3559 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3618
Db       3361 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCC 3420
```

RESULT 4
US-09-974-973-1
; Sequence 1, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3474)
US-09-974-973-1

Query Match 93.9%; Score 3398.8; DB 9; Length 3474;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy        148 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207
Db          1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
Qy        208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267
Db         61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
Qy        268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327
Db        121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
Qy        328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387
Db        181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
Qy        388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447
Db        241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
Qy        448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
Db        301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
Qy        508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567
Db        361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420
Qy        568 ACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627
Db        421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 480
Qy        628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
Db        481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540
Qy        688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
Db        541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCA 600
Qy        748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
Db        601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660
Qy        808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
Db        661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720
Qy        868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
Db        721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
Qy        928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
Db        781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840
Qy        988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
Db        841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
Qy       1048 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107
Db        901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC 960
```

```
Qy    1108 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167
Db     961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020
Qy    1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227
Db    1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
Qy    1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287
Db    1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140
Qy    1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347
Db    1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200
Qy    1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407
Db    1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
Qy    1408 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467
Db    1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320
Qy    1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527
Db    1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1380
Qy    1528 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1587
Db    1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT 1440
Qy    1588 CCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647
Db    1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
Qy    1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
Db    1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
Qy    1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
Db    1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1620
Qy    1768 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827
Db    1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
Qy    1828 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887
Db    1681 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
Qy    1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1947
Db    1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1800
Qy    1948 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 2007
Db    1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860
Qy    2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067
Db    1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920
Qy    2068 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127
Db    1921 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
Qy    2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2187
Db    1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
```

```
Qy    2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247
Db    2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
Qy    2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307
Db    2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
Qy    2308 AAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2367
Db    2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220
Qy    2368 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427
Db    2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
Qy    2428 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487
Db    2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
Qy    2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2547
Db    2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2400
Qy    2548 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607
Db    2401 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Qy    2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2667
Db    2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
Qy    2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727
Db    2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
Qy    2728 CGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC 2787
Db    2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy    2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847
Db    2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
Qy    2848 GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2907
Db    2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760
Qy    2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2967
Db    2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2820
Qy    2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 3027
Db    2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 2880
Qy    3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 3087
Db    2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 2940
Qy    3088 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3147
Db    2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000
Qy    3148 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC 3207
Db    3001 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060
Qy    3208 CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3267
Db    3061 CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
Qy    3268 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3327
```

```
Db       3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180

Qy       3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240

Qy       3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300

Qy       3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3507
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360

Qy       3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 3567
             |||||||||||||||||||||||||||||||||||||||||||||| || || |||||||||
Db       3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420

Qy       3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3474


RESULT 5
US-09-974-973-3
; Sequence 3, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
;  APPLICANT: Hanke, Paul D.
;  TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium
;  FILE REFERENCE: 1533.1230001/MAC/RGM
;  CURRENT APPLICATION NUMBER: US/09/974,973
;  CURRENT FILING DATE:  2001-10-21
;  PRIOR APPLICATION NUMBER: US 60/239,913
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3474
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-974-973-3

  Query Match             93.9%;  Score 3398.8;  DB 9;  Length 3474;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 3427;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy        148 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy        208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120

Qy        268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180

Qy        328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240

Qy        388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300

Qy        448 GTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
              ||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy        508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420

Qy        568 ACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCG 480

Qy        628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540

Qy        688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
              |||||||||||| |||||||||||||||||||||||||||||||| |||||| |||||||
Db        541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTTCA 600

Qy        748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
              |||||||||||| |||||||| ||||||||||||||||||||||||||||| ||||||||
Db        601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660

Qy        808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
              || ||  ||||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db        661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720

Qy        868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780

Qy        928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840

Qy        988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900

Qy       1048 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC 960

Qy       1108 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020

Qy       1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080

Qy       1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140

Qy       1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200

Qy       1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260

Qy       1408 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320

Qy       1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1380

Qy       1528 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1587
              |||||||||||||||||||||||||||||| || | ||||||||| ||||||||||||||||
Db       1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT 1440

Qy       1588 CCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1441 CCACCTGCGGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
Qy       1648 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707
Db       1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
Qy       1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1767
Db       1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1620
Qy       1768 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827
Db       1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
Qy       1828 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887
Db       1681 GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
Qy       1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1947
Db       1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTAC 1800
Qy       1948 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 2007
Db       1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860
Qy       2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067
Db       1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920
Qy       2068 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127
Db       1921 CCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
Qy       2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2187
Db       1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
Qy       2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247
Db       2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
Qy       2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307
Db       2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
Qy       2308 AAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2367
Db       2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220
Qy       2368 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2427
Db       2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
Qy       2428 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487
Db       2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
Qy       2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2547
Db       2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCC 2400
Qy       2548 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607
Db       2401 ATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
Qy       2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2667
Db       2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
Qy       2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727
Db       2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
```

```
Qy       2728 CGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTAC 2787
Db       2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
Qy       2788 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847
Db       2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
Qy       2848 GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2907
Db       2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760
Qy       2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2967
Db       2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAAC 2820
Qy       2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 3027
Db       2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGC 2880
Qy       3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 3087
Db       2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC 2940
Qy       3088 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3147
Db       2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000
Qy       3148 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC 3207
Db       3001 GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060
Qy       3208 CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3267
Db       3061 CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
Qy       3268 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3327
Db       3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180
Qy       3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387
Db       3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240
Qy       3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3447
Db       3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300
Qy       3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3507
Db       3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360
Qy       3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 3567
Db       3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420
Qy       3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db       3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3474
```

RESULT 6
US-10-282-122A-17701
; Sequence 17701, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17701
; LENGTH: 3423
; TYPE: DNA
; ORGANISM: Cornyebacterium diptheriae
US-10-282-122A-17701

Query Match 53.2%; Score 1927.4; DB 12; Length 3423;
Best Local Similarity 73.4%; Pred. No. 0;
Matches 2494; Conservative 0; Mismatches 896; Indels 9; Gaps 2;

```
Qy   213 ATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGT 272
Db    27 AACCTCTACGGTAAATCCATTGTCAAAGATCCTCGTTGCTAACCGTGGCGAGATCGCCGT 86
Qy   273 CCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGA 332
Db    87 GCGAGCTTTTCGTGCAGCCTTTGAAACCGGTGCTGCCACGGTAGCTGTCTACCCGAACGA 146
Qy   333 AGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGG 392
Db   147 AGACCGTAATTCATTCCACCGATCGTTTGCCTCTGAAGCTGTCCTCATTGGTGAGGGCGG 206
Qy   393 CTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAA 452
Db   207 ATCTGCGGTCAAAGCGTACCTCGATATCGACGAAATCATCCGTGCTGCTAAACAAACGGG 266
Qy   453 AGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGA 512
Db   267 TGCCGACGCAATTTACCCAGGCTACGGCTTCCTTTCCGAAAACGCCCAGCTTGCTCGTGA 326
Qy   513 GTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGG 572
Db   327 GTGCGCTGAAAACGGCATAACCTTTATCGGTCCACCTCCATCAGTGCTGGAGCTTACCGG 386
Qy   573 TGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATC 632
Db   387 TGATAAAGCAGCTGCTGTTACCGCAGCACGCGAGGCAGGTTTGCCCACGCTGACTGAGAC 446
Qy   633 CACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCAT 692
```

```
Db   447 AGAAGCAACAGACGATCCAAAGAAGCTCGCAGAGCTGACCAAGGGGCAGACCTACCCGCT 506
Qy   693 CTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGA 752
Db   507 ATTCGTCAAAGCCGTTGCCGGTGGTGGTGGCCGCGGCATGCGTTTTGTCGACGCCCCCGA 566
Qy   753 TGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGG 812
Db   567 AAACCTAGAAAAGCTAGCAGCAGAAGCCTCCCGTGAGGCGGCTGCAGCATTCGGCGATGG 626
Qy   813 CGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGG 872
Db   627 CCGAGTCTACGCCGAGCGTGCAGTAATCAATCCTCAGCATATTGAGGTCCAGATCCTTGG 686
Qy   873 CGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCA 932
Db   687 TGACTCCGCTGGCAACATCATCCACCTGTACGAGCGTGATTGCTCGCTGCAGCGTCGCCA 746
Qy   933 CCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCAT 992
Db   747 CCAAAAGGTCGTAGAGATCGCCCCAGCACAGCACTTGGCGCCCGGACTACGCGAAAAGAT 806
Qy   993 TTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGA 1052
Db   807 TTGTGCAGACGCAGTCGCGTTTGCACGCCACATCGGTTACCAAGGTGCAGGCACGGTCGA 866
Qy  1053 ATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGT 1112
Db   867 GTTCCTAGTCGATGAGGAAGGCAACCACGTCTTCATCGAGATGAACCCTCGTATCCAGGT 926
Qy  1113 TGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTT 1172
Db   927 CGAACACACCGTGACCGAAGAGGTCACCCAAGTCGACTTGGTGAAGTCCCAGATCATGAT 986
Qy  1173 GGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGG 1232
Db   987 CGCCTCCGGTGCGACACTCGAAGATCTCGGTCTGCGCCAAGAAGACATCCACACCGAAGG 1046
Qy  1233 TGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATAC 1292
Db  1047 CGCTGCCCTGCAGTGCCGTATTACTACTGAAGACCCCAACAACGGTTTCCGTCCTGACAC 1106
Qy  1293 CGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGC 1352
Db  1107 CGGAACGATTACCGCTTATCGTTCCCCAGGTGGCGCAGGTGTTCGTCTCGACGGTGCTGC 1166
Qy  1353 TCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCG 1412
Db  1167 CATGCTCGGTGGTGAGATCACCCCGAATTTCGACTCCATGCTGGTCAAGATGACCTGCCG 1226
Qy  1413 TGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGT 1472
Db  1227 TGGTGCCGATTTCGCCACCGCAGTCGCTCGCGCTCAGCGTGCACTCGCAGAATTCGTTGT 1286
Qy  1473 GTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCAC 1532
Db  1287 CTCAGGTGTTGCTACCAACATCGGATTCCTGCGTGCTTTGCTACGCGAAGAAGACTTCCA 1346
Qy  1533 TTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACC 1592
Db  1347 GCACAAGCGCATCGCCACTGGCTTCATCGCAGATCACCCATGGCTGCTGCAGGCACCACC 1406
Qy  1593 TGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCC 1652
Db  1407 TGCCGACGATGAGCCTGGCCGCATTTTGAACTACCTAGCTGATGTGACTGTGAACAAGCC 1466
Qy  1653 TCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGA 1712
Db  1467 TCATGGTCTACGCCCA---GCCGTGATCAACCCGGTAGAAAAGCTCCCAGCAGAGATCAA 1523
Qy  1713 TCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGC 1772
Db  1524 AGGTGAGCTACCACGTGGTTCCCGTGATCGCCTGCTCCAGCTCGGCCCAGAGGGATTTGC 1583
```

```
Qy       1773 TCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACA 1832
Db       1584 GCGCGCACTGCGTAAACAGGATGCGCTCGCGGTTACGGATACCACCTTCCGTGATGCACA 1643

Qy       1833 CCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGT 1892
Db       1644 TCAGTCGCTTCTCGCAACCCGCGTGCGCTCCAACACGCTGATCGACGCAGCACGGCACGT 1703

Qy       1893 CGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGT 1952
Db       1704 CGCAAAGCTCACCCCAGAGCTGCTCTCTGTAGAAGCATGGGGTGGTGCTACCTATGACGT 1763

Qy       1953 GGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGAT 2012
Db       1764 TGCTATGCGCTTCCTACATGAGGACCCCTGGGAGCGTCTCGACCACCTGCGTGAGGCCAT 1823

Qy       2013 GCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTA 2072
Db       1824 GCCCAACGTCAACATTCAGATGCTGCTGCGAGGCCGCAACACAGTCGGCTACACACCGTA 1883

Qy       2073 CCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTT 2132
Db       1884 CCCAGATTCTGTATGCAAGGCTTTTGTCGACGAAGCCGCACGCTCTGGTGTTGATATCTT 1943

Qy       2133 CCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCT 2192
Db       1944 CCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTGCT 2003

Qy       2193 GGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCC 2252
Db       2004 CGAAACCAACACCACGATTGCCGAAGTCGCAATGGCCTACTCCGGTGATCTGACCAACCC 2063

Qy       2253 AAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTC 2312
Db       2064 AAGTGAGAAGCTCTACACACTGGATTACTACCTTAAGCTTGCCGAGGAAATCGTTAAGTC 2123

Qy       2313 TGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAAC 2372
Db       2124 CGGTGCACATGTTCTAGCCATTAAAGACATGGCTGGACTGATGAAGCCAGCCGCAGCAAC 2183

Qy       2373 CAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGA 2432
Db       2184 CAAGCTGGTCACCGAACTACGACGCAACTTCGACCTGCCGGTTCATGTTCACACCCACGA 2243

Qy       2433 CACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGT 2492
Db       2244 CACTGCTGGTGGCCAGCTGGCTACGTATTGGGCTGCAGCAGCCGCTGGTGCAGATGCCGT 2303

Qy       2493 TGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGT 2552
Db       2304 CGATGGTGCCTCCGCGCCACTGTCGGGCACAACGTCGCAGCCATCGTTGTCTGCAATTGT 2363

Qy       2553 TGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCT 2612
Db       2364 GGCAGCATTCGCTAACACCTACCGCGATACGGGTCTCTCTCTGGACGCAGTCGGATCGAT 2423

Qy       2613 CGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGG 2672
Db       2424 GGAACCGTACTGGGAAGCAGTGCGCAAACTCTACGCACCCTTCGAATCCGGAACCCCAGG 2483

Qy       2673 CCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGC 2732
Db       2484 ACCAACTGGCCGCGTCTACCAGCACGAAATCCCAGGCGGCCAGCTGTCCAATCTGCGCGC 2543

Qy       2733 ACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGC 2792
Db       2544 ACAGGCAACTGCACTCGGCTTGGCAGATCGCTTCGAGCTCATCGAGGACTACTACGCAGC 2603

Qy       2793 CGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGA 2852
Db       2604 CGTTAATGAGATGCTGGGACGCCCCACCAAGGTCACCCCGTCCTCAAAGGTCGTTGGCGA 2663
```

```
Qy       2853 CCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACA 2912
Db       2664 TCTTGCACTCTACCTAGTCGGCGCAGGAGTTAACCCAGCGGACTTTGCAGCAGATCCCCA 2723

Qy       2913 AAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCC 2972
Db       2724 AAAGTACGATATTCCTGACTCCGTCATTGCCTTCCTCCGTGGTGAGCTCGGCACCCCTCC 2783

Qy       2973 AGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGC 3032
Db       2784 AGGCGGCTGGCCAGAAGAACTGCGTGCTAAGGCATTGGCAGGACGTAAGGAATCAAAGGA 2843

Qy       3033 ACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGA 3092
Db       2844 CACATTGGCCCCGCTGCCAGCAGAAGATGAAGCTCTGCTCGCAGATCGTGCGACC----- 2898

Qy       3093 ACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCA 3152
Db       2899 -GTTCGCCCAACGCTGGATCGCCTGCTCTTCCCGAAGCCTGCGGCAGAATTCGCAGAGCA 2957

Qy       3153 CCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGT 3212
Db       2958 CCGCCGTCAGTTTGGCGATACCACCAAACTCGGTGACGCAGAATTCCTCTATGGCCTCAA 3017

Qy       3213 CGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCT 3272
Db       3018 AGAAGGCAAAGAAACGGTCATTCGTACCGCAGACAGTAGCGTCCCAATGCTGGTGCGTCT 3077

Qy       3273 GGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGG 3332
Db       3078 CGATGCCGTCGGCGAGCCCGACGAAAAGGGCATGCGCAACGTTGTGTGCAATGTCAACGG 3137

Qy       3333 CCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGA 3392
Db       3138 CCAGATTCGCCCCATCCTCGTGCGTGATCGCAGCGTAGAGTCAGTGACCGCCTCGGTGGA 3197

Qy       3393 AAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGT 3452
Db       3198 AAAGGCTGATCCTTCCAACGCCGGACACGTAGCAGCACCATTCGCAGGCGTTGTCACAGT 3257

Qy       3453 GACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTAT 3512
Db       3258 GACAATCGAGGAAGGCGCTACCGTCAAAGCCGGCGATCCAGTTGCTGTCATTGAGGCAAT 3317

Qy       3513 GAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCC 3572
Db       3318 GAAGATGGAGGCAACGATCTCTGCCACCACAGACGGAACCGTCGACCGTATCGTCTTGAC 3377

Qy       3573 TGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGT 3611
Db       3378 GCAGGCCACCAAGGTCGAAGGTGGAGACTTGCTCCTCGT 3416
```

RESULT 7
US-10-282-122A-25766
; Sequence 25766, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

```
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25766
;   LENGTH: 3381
;   TYPE: DNA
;   ORGANISM: Mycobacterium avium
US-10-282-122A-25766

  Query Match             37.6%;  Score 1361.6;  DB 12;  Length 3381;
  Best Local Similarity   63.2%;  Pred. No. 0;
  Matches 2131;  Conservative    0;  Mismatches 1234;  Indels    9;  Gaps    2;

Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
              ||  || |||| || |||||||| || |||||| |||| || ||||| || ||   |||
Db         10 AAAGTCCTGGTCGCCAACCGCGGGGAGATCGCGATCCGCGCGTTCCGCGCGGCCTACGAG 69

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
              |    |||||  ||  ||  | ||  |||   || ||  ||    ||  | |||||  |
Db         70 CTGGAAATGGCCACCGTGGCGGTGTATCCCTACGAGGACCGCAATTCGGTGCACCGGTTG 129

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
              ||   ||  |   ||  ||  |||   || |||   ||||||   || ||||||   |
Db        130 AAGGCCGACGAGTCCTACCAGATCGGTGAGGAGGGCCACCCAGTCCGCGCCTACCTGTCC 189

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
              ||||  ||  ||  ||||  |||                || |||||  || |||||  ||  |||
Db        190 GTCGACGAGATCGTCGGCACGGCGCTGGCCTGCGGCGCGGATGCGATCTACCCCGGCTAC 249

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
              ||||||||||||  ||  ||  ||  ||  ||   ||| ||||   ||||||  ||  ||
Db        250 GGCTTCCTGTCGGAGAACCCCGATCTGGCGGCGGCGTGCGCGGCGGCCGGCATCACGTTC 309

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
              |  |||||  |  |  |||||  ||  ||||||||  ||  |||||  ||  ||   |  ||
Db        310 GTCGGCCCCAGCGCCGAGGTGCTCGAGCTCACCGGCGACAAGTCGCGGGCCATCGCGGCG 369

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
              |||   |   ||  |||||  ||  ||   ||||    ||   |||   |       |  |||
Db        370 GCGCGGGCCGCGGGTCTTCCGGTGCTGGCCTCCTCGCCGCCGTCGACCTCGGTGCAGGAG 429

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
              |  |  |   |||  ||      |||  |  |||   |  ||  ||  |||||  ||  |||||  ||
Db        430 TTGCTGAGCGCCGCCGAGACGATGACGTTCCCGCTGTTCGTCAAGGCGGTCGCCGGCGGC 489

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
              |||||  ||  ||  |||||    ||  |    ||  |  |||   |  ||          |
Db        490 GGCGGCCGGGGGGATGCGCCGGGTCACCGATCCGGGGGCGCTGGCCGAGGCGATCGAGGCG 549

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
              ||    |||||  ||  ||  |  ||  |||||  ||  |  |  ||||  |  |||||  |  ||  |||
Db        550 GCCAGCCGTGAGGCCGAGTCCGCGTTCGGGGACGCCTCGGTGTTTCTCGAGCAGGCGGTG 609

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
              ||  ||  ||  |  |||  ||  ||  ||  ||||||||  |  |||         ||  |  |||  |  |||
Db        610 ATCAATCCGCGGCACATCGAGGTCCAGATCCTGGCCGACACGCACGGCAACGTGATGCAC 669

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
              ||  ||  ||  ||  ||||||     |||||||  ||  |||||  ||  ||   ||||  ||  ||  ||
Db        670 CTCTACGAGCGGGACTGCAGCGTGCAGCGCCGGCACCAGAAGGTGATCGAGATCGCCCCG 729

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
              ||  |   |  |||||||  |   |||||  ||  ||  ||  ||  ||  ||  ||    ||||
Db        730 GCGCCCAACCTGGATCCGGCGCTGCGCGAACGTATCTGCGCCGACGCGGTGGCGTTCGCA 789

Qy        1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
               |||  |||  ||  |||     |||||||  ||  |||||  |||  ||  ||||  |||    ||||||
Db         790 CGCAGCATCGGCTACACCTGCGCGGGCACTGTGGAGTTCCTGCTCGACGAACGCGGCAAC 849

Qy        1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
               |||||  ||||||||  |||||||||  ||  ||||||||  ||||||||  |||||  ||  ||   ||
Db         850 CACGTGTTCATCGAGATGAACCCGCGCATCCAGGTGGAGCACACGGTGACCGAGGAGATC 909

Qy        1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
               |||||  ||  |||||||||     ||||||  |||||  |  ||   |  ||     ||  ||  |  ||
Db         910 ACCGACGTCGACCTGGTGTCCGCGCAGCTGCGCATCGCCTCCGGCCAGACGCTGGAAGAG 969

Qy        1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
               |  ||  ||||  |||  |||   |  |||   |  |  |||  ||  ||  ||||||||||  ||||||
Db         970 ATTGGCCTGAGCCAGGATTCGGTCATCTCACGCGGCGCGGCCCTGCAGTGCCGGATCACC 1029

Qy        1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
               ||  ||  ||  ||    ||||||  ||||||||  ||  |||||    ||||||||  ||||||  |
Db         1030 ACCGAGGACCCCGCCAACGGTTTCCGCCCGGACACCGGCCGCATCACCGCCTACCGCACC 1089

Qy        1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
               ||  ||  ||  ||  |||  |  ||  ||  |||||  |    |    |||  ||  |  ||  ||||  |||
Db         1090 CCGGGTGGCGCGGGCATCCGGCTGGACGGCGGCACCACGCTGGGCGCGGAGATCAGCGCG 1149

Qy        1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
               |||||  ||  ||  ||||||  |  ||   ||||||||||  ||   ||||||     ||  ||  ||
Db         1150 CACTTCGATTCGATGCTGATCAAGCTGACCTGCCGGGGACGCGACTTCCCCACGGCGGTG 1209

Qy        1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
                  ||  ||  ||||||  |  ||  ||  |||   |  |    ||  ||   |  ||  ||  ||
Db         1210 CGGCGGGCGCGGCGCGCGGTCGCCGAATTCCGCATCCGCGGCGTCTCGACGAATATCCCG 1269

Qy        1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
               |||  |||  |||  |  |||    ||    |||  ||       |  |||||||  |||||    |||
Db         1270 TTCCTGCAAGCGGTTCTGGACGACCCGGATTTTCAGGCTGGGCGCATCACCACCTCGTTC 1329

Qy        1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
               ||  |   |  |||||  |  ||    ||  |    |  |  |  ||                |||
Db         1330 ATCGAACAGCGGCCGCAGTTGCTCACCGCGCGCAGTTCGGCCGACCGCGGCACCAAGATC 1389

Qy        1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
               ||  |  |||  ||||  ||  ||  |||||  ||||||||  ||  ||  |  |||  ||    ||  ||
Db         1390 CTCAACTACCTGGCCGACGTGACCGTCAACAAGCCGCACGGCGAGCGCCCGTCGGCGGTG 1449

Qy        1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
                  |  |  |  |  |  ||     |  |  |  ||     |  |  |||     ||  |||
Db         1450 TACCCGCATGACAAGCTGCCCGACATCGACCTGGCGGCGAGCCCGCCGGCGGGCTCCAAG 1509

Qy        1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
               |  ||  |||  |  ||  ||  ||  |    ||  ||          ||  ||  |||  ||  ||
Db         1510 CAGCGGTTGACCGAACTCGGGCCGGAGCGATTCGCGGCCTGGCTGCGCGAGTCGGGCGGC 1569

Qy        1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
               |  |  ||  ||  |||||  ||||||||  ||  |||||||  ||||  ||||||||  |||
Db         1570 GTCGGCGTCACCGACACCACGTTCCGCGACGCGCACCAGTCGCTGCTGGCGACCCGGGTC 1629

Qy        1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
               |||  |  ||  |||  |  ||||      |||||  |  |  ||  ||||  |||
Db         1630 CGCACCAGCGGGCTGCTCAAGGTGGCACCCTACATCGCACGGACCATGCCGCAGCTGTTG 1689
```

```
Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db       1690 TCGGTGGAGTGTTGGGGCGGAGCGACTTACGACGTGGCGCTGCGGTTTCTCAAGGAGGAC 1749

Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db       1750 CCGTGGGAGCGGCTGACGGCGCTGCGCGAGGCGATGCCGAACATCTGCCTTCAGATGCTG 1809

Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db       1810 CTGCGTGGGCGTAACACCGTGGGGTACACGCCCTATCCGGAGACGGTCACGACGGCCTTC 1869

Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db       1870 GTCGCGGAGGCGACCGAGACCGGCATCGACATCTTCCGCATCTTCGACGCGTTGAACAAC 1929

Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db       1930 GTCGACTCGATGCGCCCGGCCATCGACGCGGTCCGCGAAACCGGAAGCGCCATAGCCGAA 1989

Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db       1990 GTGGCGATGTCCTACACCGGCGACCTGTCCGATCCGGCCGAAAAGCTTTACACGCTGGAC 2049

Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db       2050 TATTACCTGAAGCTGGCCGACCAGATCGTCGAAGCGGGCGCGCACGTGCTGGCGATCAAG 2109

Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db       2110 GACATGGCCGGCCTGCTGCGCGCACCGGCGGCCGCGACGCTGGTGTCGGCGCTGAAGTCG 2169

Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db       2170 CGGTTCGACCTGCCGGTGCACGTGCACACCCATGACACGCCGGGCGGACAGCTGGCGACC 2229

Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db       2230 TACGTGGCGGCCTGGCAGGCCGGGGCCGATGCGGTCGACGGCGCCGCCGCGCCGCTCGCC 2289

Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db       2290 GGAACCACCAGCCAGCCGGCACTATCGTCGATCGTGGCCGCCGCGGCCAACACGGAGTAT 2349

Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db       2350 GACACCGGGTTGTCCCTGCCGGCGGTGTGCGATCTGGAGCCGTACTGGGAGGCGCTGCGA 2409

Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db       2410 AAGGTATACGCTCCCTTCGAATCCGGTCTTCCGGCGCCGACCGGGCGGGTGTATCACCAC 2469

Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db       2470 GAGATCCCGGGCGGCCAATTGTCGAATCTGCGTCAGCAGGCGATTGCGCTGGGACTCGGC 2529

Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db       2530 GACCGGTTCGAGGACATCGAAAACGCCTACGCCGGCGCCGACGCGATTTTGGGACACCTG 2589

Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db       2590 GTCAAGGTCACACCGTCCAGCAAGGTGGTGGGCGACCTGGCGCTGGCCCTGGTGGGTGCG 2649

Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db       2650 GGTGTGGGCGCGCAGGACTTCGCCGAAGACCCCTCCCGCTACGACATCCCGGATTCGGTG 2709

Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db       2710 ATCGGCTTTCTGCGCGGCGAGCTCGGTGACCCGCCGGGCGGCTGGCCGGAGCCGTTGCGC 2769

Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057

Db       2770 ACCAAGGCCTTACAGGGCCGGGGGCCGGCCAAGCCGGAGCAGCCGCTGACGGCCGAGGAC 2829

Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db       2830 GAAGCGGCGTTGGCCGCCCCGGGTG------CGCGACGGCAGGCAGCGCTGAACCGGCTG 2883

Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db       2884 CTGTTCCCGGGACCGACCAAGGAGCTCGAGGAGCACCGCGAGCAGTACGGTGACACCTCG 2943

Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db       2944 GGGCTGAGCGCCAACCAGTTCTTCTACGGATTGCGGCAGGGCGACGAACACCGCGTCGAA 3003

Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db       3004 CT---CGAGCGCGGCGTGGAGTTGCTGATCGGCCTGGAGGCCATCTCCGATCCCGACGAA 3060

Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db       3061 CGCGGCATGCGCACCGTGATGTGCATTCTCAACGGGCAGCTGCGGCCGGTGGTGGTGCGC 3120

Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db       3121 GACCGCAGCATCGCCATCGACGTGCCCGCCGCCGAGAAGGCCGACCGCGCCAACCCGGAC 3180

Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
Db       3181 CACATCGCTGCCCCGTTCGCCGGGGTGGTGACGGTCGCCGCCGAGGTCGGCCGGCAGGTC 3240

Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
Db       3241 GAGGCCGGCCAGACCATCGCGACGATCGAGGCGATGAAGATGGAAGCCGCGGTCACCAGC 3300

Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
Db       3301 CCGAAGTCGGGCAAGGTGGCGCGAATCGCCGTGTCGCGGACCGCTCAAGTCGAGGGCGGC 3360

Qy       3598 GACTTGATCGTCGT 3611
Db       3361 GACCTCCTGATGGT 3374
```

RESULT 8
US-10-282-122A-28626
; Sequence 28626, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28626
;   LENGTH: 3384
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28626

  Query Match             35.7%;  Score 1294.4;  DB 12;  Length 3384;
  Best Local Similarity   62.0%;  Pred. No. 0;
  Matches 2099;  Conservative    0;  Mismatches 1276;  Indels    9;  Gaps    3;

Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db         10 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 69

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db         70 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 129

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db        130 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 189

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db        190 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 249

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db        250 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 309

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db        310 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 369

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db        370 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 429

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db        430 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 489

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db        490 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 549

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db        550 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 609

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db        610 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 669

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db        670 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 729

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db        730 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 789

Qy       1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db        790 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 849

Qy       1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db        850 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 909

Qy       1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db        910 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 969

Qy       1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db        970 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 1029

Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db       1030 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 1089

Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db       1090 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 1149

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db       1150 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 1209

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db       1210 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 1269

Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db       1270 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 1329

Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db       1330 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 1389

Qy       1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db       1390 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 1449

Qy       1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db       1450 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 1509

Qy       1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db       1510 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 1569

Qy       1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db       1570 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 1629

Qy       1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db       1630 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 1689

Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db       1690 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 1749

Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db       1750 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 1809

Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
```

```
Db       1810 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 1869
Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db       1870 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 1929
Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
Db       1930 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 1989
Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db       1990 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 2049
Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db       2050 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 2109
Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db       2110 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 2169
Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db       2170 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 2229
Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db       2230 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCGCCGCTGGCG 2289
Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db       2290 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 2349
Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db       2350 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 2409
Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db       2410 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 2469
Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db       2470 GAGATTCCGGGCGGCCAACTGTCCAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 2529
Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db       2530 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 2589
Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db       2590 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 2649
Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db       2650 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 2709
Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db       2710 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 2769
Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db       2770 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 2826
Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db       2827 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 2883
Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
```

```
Db       2884 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 2943
Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db       2944 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 3003
Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db       3004 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 3060
Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db       3061 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 3120
Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db       3121 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 3180
Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
Db       3181 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTGCGTCGGCGAGCGGGTC 3240
Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
Db       3241 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 3300
Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
Db       3301 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 3360
Qy       3598 GACTTGATCGTCGTCGTTTCCTAA 3621
Db       3361 GACCTGTTGGTGGTGGTGAGCTGA 3384
```

RESULT 9
US-10-282-122A-26500
; Sequence 26500, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636

```
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26500
;   LENGTH: 3381
;   TYPE: DNA
;   ORGANISM: Mycobacterium bovis
US-10-282-122A-26500

  Query Match             35.7%;  Score 1292.2;  DB 12;  Length 3381;
  Best Local Similarity   62.0%;  Pred. No. 0;
  Matches 2095;  Conservative    0;  Mismatches 1273;  Indels    9;  Gaps    3;

Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db         10 AAGGTGCTCGTCGCCAATCGCGGGGAGATCGCGATCCGGGCCTTTCGTGCCGCCTACGAA 69

Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db         70 CTGGGCGTCGGAACCGTGGCCGTTTATCCGTACGAGGACCGCAATTCGCAGCACCGTCTC 129

Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db        130 AAGGCGGACGAGTCTTACCAGATCGGCGACATCGGTCACCCGGTGCATGCATACCTGTCG 189

Qy        418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db        190 GTCGACGAGATCGTCGCGACGGCCCGTCGGGCGGGTGCCGACGCTATCTACCCTGGCTAC 249

Qy        478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db        250 GGGTTTCTATCGGAGAATCCGGATCTGGCTGCGGCATGCGCGGCGGCGGGCATCAGCTTC 309

Qy        538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db        310 GTCGGTCCCAGCGCCGAAGTGCTTGAGCTGGCTGGGAATAAGTCTCGCGCCATCGCGGCG 369

Qy        598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAG 657
Db        370 GCCCGCGAAGCCGGCTTGCCCGTGCTGATGTCCTCGGCGCCGTCGGCCTCGGTCGACGAA 429

Qy        658 ATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGT 717
Db        430 CTGCTGTCGGTTGCGGCCGGCATGCCGTTTCCGTTGTTCGTCAAGGCAGTTGCCGGTGGC 489

Qy        718 GGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAA 777
Db        490 GGGGGCCGGGGTATGCGTCGTGTCGGCGATATCGCGGCGCTTCCGGAGGCGATCGAAGCC 549

Qy        778 GCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTG 837
Db        550 GCCAGCCGGGAAGCCGAGTCGGCGTTCGGGGACCCGACGGTCTATCTCGAGCAGGCAGTG 609

Qy        838 ATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACAC 897
Db        610 ATCAATCCACGCCACATCGAGGTGCAGATTCTGGCGGACAACCTCGGCGACGTGATCCAT 669

Qy        898 CTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCA 957
Db        670 CTCTATGAGCGTGACTGCAGTGTGCAGCGTCGCCATCAGAAGGTCATCGAGCTGGCGCCC 729

Qy        958 GCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGC 1017
Db        730 GCGCCGCACCTGGACGCCGAGTTGCGTTACAAGATGTGCGTCGATGCGGTCGCCTTCGCC 789

Qy       1018 CGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAAC 1077
Db        790 CGCCATATCGGGTACAGCTGCGCGGGCACCGTCGAGTTCCTGCTGGACGAGCGAGGGGAG 849

Qy       1078 CACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTC 1137
Db        850 TATGTCTTCATCGAGATGAATCCGCGGGTTCAGGTGGAGCACACGGTGACCGAGGAGATT 909

Qy       1138 ACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAA 1197
Db        910 ACCGACGTCGACCTGGTCGCCAGCCAGCTGCGCATTGCCGCCGGGGAGACGCTCGAACAA 969

Qy       1198 TTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACC 1257
Db        970 TTGGGCCTGCGGCAGGAGGACATCGCACCGCATGGCGCCGCACTACAGTGCCGGATCACC 1029

Qy       1258 ACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCA 1317
Db       1030 ACCGAGGATCCGGCCAACGGCTTCCGGCCGGACACGGGCCGGATCAGCGCGTTGCGCACC 1089

Qy       1318 CCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
Db       1090 GCCGGCGGTGCCGGTGTCCGCCTGGACGGCAGCACCAACCTGGGCGCAGAAATCAGCCCG 1149

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
Db       1150 TACTTCGACTCCATGCTGGTCAAGCTGACCTGTCGGGGCCGTGACCTCCCTACCGCAGTG 1209

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
Db       1210 AGCCGTGCGCGCCGGGCGATCGCGGAGTTCCGGATCCGCGGGGTATCGACGAATATTCCG 1269

Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
Db       1270 TTCCTGCAAGCGGTCCTGGATGACCCGGACTTCCGAGCGGGCCGGGTCACCACGTCCTTC 1329

Qy       1558 ATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATC 1617
Db       1330 ATTGATGAGCGGCCGCAGCTGCTGACCGCGCGCGCCTCGGCCGACCGCGGCACCAAGATC 1389

Qy       1618 CTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTT 1677
Db       1390 CTTAACTTCCTGGCCGATGTCACCGTCAACAACCCGTATGGCTCGCGTCCGTCAACGATC 1449

Qy       1678 GCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGT 1737
Db       1450 TACCCGGACGACAAGCTGCCCGATCTTGATCTGCGGGCCGCACCACCGGCCGGGTCCAAG 1509

Qy       1738 GACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCA 1797
Db       1510 CAGCGACTAGTCAAGTTGGGGCCGGAAGGATTTGCTCGTTGGCTGCGGGAGTCGGCCGCG 1569

Qy       1798 CTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTC 1857
Db       1570 GTCGGGGTCACCGATACCACATTCCGGGATGCTCACCAGTCGTTACTGGCTACCCGAGTA 1629

Qy       1858 CGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTG 1917
Db       1630 CGCACCAGCGGACTGTCGCGGGTGGCACCGTATCTCGCGCGGACCATGCCGCAGCTGTTG 1689

Qy       1918 TCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGAT 1977
Db       1690 TCCGTGGAGTGTTGGGGCGGTGCGACTTACGATGTGGCGCTGCGCTTTCTCAAGGAGGAT 1749

Qy       1978 CCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTG 2037
Db       1750 CCCTGGGAACGGCTGGCCACACTGCGTGCAGCAATGCCCAATATCTGCTTGCAGATGCTA 1809

Qy       2038 CTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTT 2097
Db       1810 TTGCGGGGCCGCAATACCGTGGGCTACACGCCGTACCCGGAAATCGTGACGTCGGCGTTT 1869

Qy       2098 GTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGAC 2157
Db       1870 GTGCAAGAAGCAACAGCCACCGGTATCGACATCTTTCGTATCTTCGACGCGCTCAACAAC 1929

Qy       2158 GTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAG 2217
```

```
Db       1930 ATCGAGTCGATGCGTCCGGCGATCGACGCAGTACGCGAAACAGGTTCTGCGATAGCAGAA 1989
Qy       2218 GTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGAT 2277
Db       1990 GTCGCGATGTGCTACACGGGCGACCTTACCGATCCGGGTGAACAGCTGTACACGCTGGAC 2049
Qy       2278 TACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAG 2337
Db       2050 TACTACCTGAAACTGGCTGAGCAGATCGTGGACGCCGGCGCCCATGTGCTGGCGATCAAG 2109
Qy       2338 GATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGT 2397
Db       2110 GATATGGCTGGACTGTTGCGCCCGCCGGCCGCTCAAAGGTTGGTCAGCGCCCTGCGCAGT 2169
Qy       2398 GAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACC 2457
Db       2170 CGCTTCGACCTGCCCGTTCACCTGCACACCCACGACACACCGGGTGGCCAGCTCGCCAGC 2229
Qy       2458 TACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCT 2517
Db       2230 TATGTGGCCGCTTGGCACGCCGGGGCCGATGCCGTCGACGGCGCCGCCGCGCCGCTGGCG 2289
Qy       2518 GGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGC 2577
Db       2290 GGAACGACCAGCCAGCCCGCGCTGAGCTCGATCGTTGCTGCCGCTGCCCACACCGAGTAC 2349
Qy       2578 GATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGC 2637
Db       2350 GACACCGGCCTGTCGCTTTCGGCGGTGTGCGCCCTGGAGCCGTACTGGGAGGCGTTACGA 2409
Qy       2638 GGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCAC 2697
Db       2410 AAAGTGTATGCGCCGTTCGAGTCTGGGTTGCCGGGGCCGACGGGGCGGGTTTATCACCAC 2469
Qy       2698 GAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCG 2757
Db       2470 GAGATTCCGGGCGGCCAACTGTCTAATCTGCGCCAGCAAGCAATTGCTCTTGGTCTGGGA 2529
Qy       2758 GATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCA 2817
Db       2530 GATCGATTCGAAGAGATCGAAGAGGCCTACGCGGGCGCCGACCGAGTGTTGGGCAGGCTG 2589
Qy       2818 ACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCG 2877
Db       2590 GTTAAGGTCACGCCAACGTCCAAGGTGGTCGGCGATCTGGCGCTGGCACTGGTCGGCGCC 2649
Qy       2878 GGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC 2937
Db       2650 GGTGTCAGTGCAGACGAATTCGCCTCCGATCCAGCGCGATTTGGCATCCCGGAATCGGTA 2709
Qy       2938 ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGC 2997
Db       2710 CTCGGATTTCTGCGGGGCGAGCTGGGTGATCCGCCCGGCGGGTGGCCCGAACCGCTGCGC 2769
Qy       2998 ACCCGCGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAA 3057
Db       2770 ACTGCGGCGCTGGCCGGTCGCGGGGCGGCCAGGCCCACT---GCGCAATTGGCGGCGGAC 2826
Qy       3058 GAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCCTG 3117
Db       2827 GACGAGATTGCCCTATCGTCGGTCGGAGCCAAG---CGTCAGGCCACCCTGAACAGGCTG 2883
Qy       3118 CTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCT 3177
Db       2884 TTATTTCCCAGTCCAACAAAGGAATTCAATGAGCACCGGGAAGCCTACGGCGACACGTCG 2943
Qy       3178 GCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATCCGC 3237
Db       2944 CAATTGTCGGCCAACCAGTTCTTCTATGGTCTGCGCCAAGGTGAAGAGCATCGGGTGAAG 3003
Qy       3238 CTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGACGAT 3297
Db       3004 CTG---GAGCGTGGGGTGGAGCTGTTGATCGGGCTGGAGGCCATTTCCGAACCCGACGAA 3060

Qy       3298 AAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGT 3357
Db       3061 CGCGGCATGCGAACGGTGATGTGCATCCTCAACGGGCAGCTGCGGCCGGTGCTAGTGCGC 3120
Qy       3358 GACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGC 3417
Db       3121 GACCGCAGCATTGCCAGTGCCGTTCCGGCCGCCGAGAAGGCCGACCGCGGCAATCCCGGA 3180
Qy       3418 CATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGATGAGGTC 3477
Db       3181 CACATCGCCGCGCCATTTGCCGGAGTCGTCACGGTCGGGGTGTGCGTCGGCGAGCGGGTC 3240
Qy       3478 AAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCT 3537
Db       3241 GGCGCCGGCCAAACCATCGCCACCATCGAGGCGATGAAGATGGAAGCCCCGATCACCGCC 3300
Qy       3538 TCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGC 3597
Db       3301 CCGGTTGCCGGCACCGTGGAGCGGGTGGCGGTGTCGGACACCGCCCAGGTGGAGGGCGGA 3360
Qy       3598 GACTTGATCGTCGTCGT 3614
Db       3361 GACCTGTTGGTGGTGGT 3377
```

```
RESULT 10
US-09-917-800A-1566
; Sequence 1566, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1566
;   LENGTH: 3945
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012744
US-09-917-800A-1566

  Query Match             18.8%;  Score 681;  DB 9;  Length 3945;
  Best Local Similarity   54.2%;  Pred. No. 3.2e-197;
  Matches 1521;  Conservative    0;  Mismatches 1260;  Indels   24;  Gaps    6;
```

```
Qy       233 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
Db       144 TCAAGAAAGTAATGGTGGCCAACAGAGGTGAGATTGCCATCCGAGTGTTTCGTGCCTGCA 203

Qy       293 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 352
Db       204 CAGAGCTGGGTATCCGCACAGTGGCTGTCTACTCGGAGCAGGACACAGGCCAGATGCACC 263

Qy       353 GCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACC 412
Db       264 GGCAGAAAGCTGATGAAGCCTACCTTATTGGCCGTGGGCTGGCTCCTGTGCAAGCCTACC 323

Qy       413 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGG 472
Db       324 TGCACATTCCAGACATCATTAAGGTGGCCAAGGAGAATGGTGTAGATGCTGTGCACCCTG 383

Qy       473 GATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTA 532
Db       384 GCTATGGGTTCCTCTCAGAGAGAGCAGACTTTGCCCAGGCCTGCCAAGATGCTGGAGTCC 443

Qy       533 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAA 592
Db       444 GATTCATTGGTCCAAGCCCAGAGGTGGTCCGCAAGATGGGAGACAAGGTGGAAGCCCGGG 503

Qy       593 CCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCA---CCCCGAGCAAAAACA 649
Db       504 CCATTGCCATTGCTGCAGGCGTTCCAGTGGTCCCTGGCACTAATTCCCCCATCAATTCCC 563

Qy       650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
Db       564 TGCATGAGGCACACGAGTTCTCTAACACCTATGGTTTCCCTATTATCTTCAAGGCTGCCT 623

Qy       710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
Db       624 ATGGAGGTGGGGGCCGTGGCATGAGGGTTGTGCATAGCTACGAGGAGCTGGAAGAGAATT 683

Qy       770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
Db       684 ACACCCGGGCCTACCCTGAGGCCTTGGCAGCCTTTGGGAATGGGGCATTGTTTGTGGAGA 743

Qy       830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
Db       744 AATTCATTGAGAAGCCAAGACACATTGAGGTGCAGATCCTAGGGGACCAATATGGGAACA 803

Qy       890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
Db       804 TCTTGCACTTGTATGAGCGGGACTGCTCCATCCAGCGGCGGCACCAGAAGGTGGTAGAGA 863

Qy       950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
Db       864 TTGCCCCTGCTACCCACCTGGACCCCCAACTTCGGTCACGCCTCACCAGTGACTCTGTCA 923

Qy      1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
Db       924 AACTTGCCAAGCAGGTTGGCTATGAGAATGCAGGCACTGTGGAGTTCCTGGTGGACAAGC 983

Qy      1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
Db       984 ATGGCAAGCACTACTTCATCGAGGTCAATTCCCGCCTGCAGGTGGAGCACACGGTCACTG 1043

Qy      1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
Db       1044 AGGAGATTACAGATGTGGACCTGGTCCATGCTCAGATCCATGTGTCCGAAGGCCGGAGCC 1103

Qy      1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
Db       1104 TGCCTGACCTAGGCCTGCGGCAGGAAAACATCCGAATCAATGGTTGTGCCATTCAGTGTC 1163

Qy      1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
Db       1164 GGGTCACCACTGAGGACCCTGCACGCAGCTTCCAGCCAGACACTGGCCGCATTGAGGTTT 1223

Qy      1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCG 1366
```

```
Db       1224 TCCGGAGTGGTGAGGGCATGGGCATCCGCCTGGACAATGCCTCAGCATTCCAGGGAGCTG 1283

Qy      1367 AAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTG 1426
Db       1284 TCATATCCCCCCACTATGACTCCCTGCTCGTCAAAGTCATTGCCCATGGCAAAGACCACC 1343

Qy      1427 AAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAA 1486
Db       1344 CTACAGCTGCCACCAAGATGAGCAGAGCCCTGGCGGAGTTCCGTGTCCGAGGTGTAAAGA 1403

Qy      1487 CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCG 1546
Db       1404 CCAACATCCCCTTCCTGCAGAATGTGCTCAACAACCAGCAGTTCCTAGCGGGCATTGTGG 1463

Qy      1547 CCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGC 1606
Db       1464 ACACCCAGTTCATCGATGAGAACCCCGAGCTGTTCCAGCTGCGGCCTGCACAGAACCGGG 1523

Qy      1607 AGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTC 1666
Db       1524 CCCAGAAGTTGCTACATTACCTTGGACACGTCATGGTCAATGGCCCTACCACTCCAATCC 1583

Qy      1667 CAAAGGATGTTGCAGCTCCTATCGATAAGC------TGCCTAACATCAAGGATCTGCCAC 1720
Db       1584 CCGTCAAGGTCAGTCCCAGCCCTGTGGACCCCATTGTTCCTGTGGTGCCCATAGGCCCAC 1643

Qy      1721 TGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC 1780
Db       1644 CCCCAGCTGGTTTCAGAGACATCCTTCTGCGAGAGGGGCCAGAGGGCTTTGCCAGAGCTG 1703

Qy      1781 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTT 1840
Db       1704 TGCGGAATCACCAGGGGCTGCTGCTAATGGACACAACCTTCCGGGATGCCCACCAGTCAC 1763

Qy      1841 TGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGC 1900
Db       1764 TACTTGCCACTAGAGTGCGCACACACGATCTCAAAAAGATTGCACCCTACGTTGCCCACA 1823

Qy      1901 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGC 1960
Db       1824 ACTTCAACAACCTCTTCAGCATAGAGAACTGGGGAGGAGCCACATTTGACGTGGCCATGC 1883

Qy      1961 GTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 2020
Db       1884 GCTTCTTGTATGAGTGCCCCTGGCGGCGGCTCCAGGAGCTCCGGGAGCTCATCCCCAACA 1943

Qy      2021 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080
Db       1944 TCCCATTCCAGATGCTACTGAGGGGGGCCAATGCTGTGGGCTACACCAACTACCCTGACA 2003

Qy      2081 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 2140
Db       2004 ACGTGGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGAATGGCATGGACGTCTTCCGGATCT 2063

Qy      2141 TCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCA 2200
Db       2064 TTGACTCCCTTAACTACCTGCCAAACATGCTGCTGGGCATGGAAGCAG---CTGGCAGTG 2120

Qy      2201 ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAA 2260
Db       2121 CTGGGGGTGTGGTGGAAGCTGCCATCTCCTACACGGGTGACGTGGCTGACCCCAGTCGCA 2180

Qy      2261 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2320
Db       2181 CTAAATACTCACTGGAGTACTACATGGGCTTAGCTGAAGAACTGGTGCGAGCCGGCACTC 2240

Qy      2321 ACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2380
Db       2241 ACATCCTCTGCATTAAGGACATGGCAGGCCTGCTGAAGCCTGCAGCATGCACCATGCTGG 2300

Qy      2381 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTG 2437
```

```
Db       2301 TCAGCTCCCTCCGGGACCGGTTCCCCGACCTCCCACTGCACATCCATACCCATGACACAT 2360
Qy       2438 CGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497
Db       2361 CAGGGTCAGGTGTGGCAGCCATGTTGGCCTGTGCACAAGCTGGGGCTGATGTTGTGGATG 2420
Qy       2498 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTG 2557
Db       2421 TGGCAGTCGACTCTATGTCTGGGATGACCTCACAGCCCAGCATGGGGGCCCTGGTGGCCT 2480
Qy       2558 CATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGC 2617
Db       2481 GTACCAAAGGGACTCCTCTGGACACAGAGGTACCCCTGGAGCGTGTGTTTGACTACAGTG 2540
Qy       2618 CGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGA------GTCTGGAACCCCAG 2671
Db       2541 AGTATTGGGAAGGGGCTCGGGGGCTGTATGCAGCCTTTGATTGCACGGCTACCATGAAGT 2600
Qy       2672 GCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTG 2731
Db       2601 CTGGCAACTCAGACGTGTATGAGAATGAGGATCCAGGGGGCCAGTACACCAACCTACACT 2660
Qy       2732 CACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAG 2791
Db       2661 TCCAGGCCCACAGCATGGGACTTGGCTCCAAGTTCAAGGAGGTCAAGAAGGCCTATGTGG 2720
Qy       2792 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCG 2851
Db       2721 AGGCTAACCAGATGCTGGGGGACCTCATCAAGGTGACACCATCCTCCAAGATTGTGGGGG 2780
Qy       2852 ACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCAC 2911
Db       2781 ATCTGGCCCAGTTCATGGTGCAGAACGGGTTGAGCCGGGCAGAGGCAGAAGCTCAGGCAG 2840
Qy       2912 AAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTC 2971
Db       2841 AAGAGCTGTCCTTCCCCCGCTCTGTGGTGGAGTTCCTGCAGGGCTACATTGGCATTCCCC 2900
Qy       2972 CAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCC 3016
Db       2901 ATGGGGGTTTCCCTGAACCCTTCCGTTCTAAGGTGCTAAAGGACC 2945
```

RESULT 11
US-09-880-107-3029
; Sequence 3029, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3029
; LENGTH: 4017
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 S72370
US-09-880-107-3029

Query Match 18.7%; Score 677.4; DB 9; Length 4017;
Best Local Similarity 54.4%; Pred. No. 4.1e-196;
Matches 1525; Conservative 0; Mismatches 1256; Indels 24; Gaps 7;

```
Qy        233 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
Db        192 TCAAGAAAGTCATGGTGGCCAACAGAGGTGAGATTGCCATCCGTGTGTTCCGGGCCTGCA 251
Qy        293 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 352
Db        252 CGGAGCTGGGCATCCGCACCGTAGCCATCTACTCTGAGCAGGACACGGGCCAGATGCACC 311
Qy        353 GCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACC 412
Db        312 GGCAGAAAGCAGATGAAGCCTATCTCATCGGCCGCGGCCTGGCCCCCGTGCAGGCCTACC 371
Qy        413 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGG 472
Db        372 TGCACATCCCAGACATCATCAAGGTGGCCAAGGAGAACAACGTAGATGCAGTGCACCCTG 431
Qy        473 GATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTA 532
Db        432 GCTACGGGTTCCTCTCTGAGCGAGCGGACTTCGCCCAGGCCTGCCAGGATGCAGGGGTCC 491
Qy        533 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAA 592
Db        492 GGTTTATTGGTCCAAGCCCAGAAGTGGTCCGCAAGATGGGAGACAAGGTGGAGGCCCGGG 551
Qy        593 CCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGG---AATCCACCCCGAGCAAAAACA 649
Db        552 CCATCGCCATTGCTGCGGGTGTTCCCGTGGTCCCTGGCACAGATGCCCCCATCACGTCCC 611
Qy        650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
Db        612 TGCATGAGGCCCACGAGTTCTCCAACACCTACGGCTTCCCCATCATCTTCAAGGCGGCCT 671
Qy        710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
Db        672 ATGGGGGTGGGGGGCGTGGCATGAGGGTGGTGCACAGCTACGAGGAGCTGGAGGAGAATT 731
Qy        770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
Db        732 ACACCCGGGCCTACTCAGAGGCCTGGCCGGCCTTTGGGAATGGGGCGCTGTTTGTGGAGA 791
Qy        830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
Db        792 AGTTCATCGAGAAGCCACGGCACATCGAGGTGCAGATCTTGGGGGACCAGTATGGGAACA 851
Qy        890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
Db        852 TCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGGCGGCACCAGAAGGTGGTCGAGA 911
Qy        950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
Db        912 TTGCCCCCGCCGCCCACCTGGACCCGCAGCTTCGGACTCGGCTCACCAGCGACTCTGTGA 971
Qy       1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
Db        972 AACTCGCTAAACAGGTGGGCTACGAGAACGCAGGCACCGTGGAGTTCCTGGTGGACAGGC 1031
Qy       1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
Db       1032 ACGGCAAGCACTACTTCATCGAGGTCAACTCCCGCCTGCAGGTGGAGCACACGGTCACAG 1091
Qy       1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
Db       1092 AGGAGATCACCGACGTAGACCTGGTCCATGCTCAGATCCACGTGGCTGAAGGCAGGAGCC 1151
Qy       1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
Db       1152 TACCCGACCTGGGGCTGCGGCAGGAGAACATCCGCATCAACGGGTGTGCCATCCAGTGCC 1211
Qy       1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
Db       1212 GGGTCACCACCGAGGACCCCGCGCCCACGTTCCAGCCGGACACCGGCCGCATTGAGGTGT 1271
```

```
Qy       1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1366
Db       1272 TCCGGAGCGGAGAGGGCATGGGCATCCGCCTGGATAATGCTTCCGCCTTCCAAGGAGCCG 1331
Qy       1367 AAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTG 1426
Db       1332 TCATCTCGCCCCACTACGACTCCCTGCTGGTCAAAGTCATTGCCCACGGCAAAGACCACC 1391
Qy       1427 AAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAA 1486
Db       1392 CCACGGCCGCCACCAAGATGAGCAGGGCCCTTGCGGAGTTCCGTGTCCGAGGTGTGAAGA 1451
Qy       1487 CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCG 1546
Db       1452 CCAACATCGCCTTCCTGCAGAATGTGCTCAACAACCAGCAGTTCCTGGCAGGCACTGTGG 1511
Qy       1547 CCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGC 1606
Db       1512 ACACCCAGTTCATCGACGAGAACCCAGACGTGTTCCAGCTGCGGCCTGCACAGAACCGGG 1571
Qy       1607 AGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAA----CAAGCCTCATGGTGTG 1662
Db       1572 CCCAAAAGCTGTTGCACTACCTCGGCCATGTCATGGTAAACGGTCCAACCACCCCGATTC 1631
Qy       1663 CGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG--CCTAACATCAAGGATCTGCCAC 1720
Db       1632 CCGTCAAGGCCAGCCCCAGCCCCACGGACCCCGTTGTCCCTGCAGTGCCCATAGGCCCGC 1691
Qy       1721 TGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC 1780
Db       1692 CCCCGGCTGGTTTCAGAGACATCCTGCTGCGAGAGGGGCCTGAGGGCTTTGCTCGAGCTG 1751
Qy       1781 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTT 1840
Db       1752 TGCGGAACCACCCGGGGCTGCTGCTGATGGACACGACCTTCAGGGACGCCCACCAGTCAC 1811
Qy       1841 TGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGC 1900
Db       1812 TGCTGGCCACTCGTGTGCGCACCCACGATCTCAAAAAGATCGCCCCCTATGTTGCCCACA 1871
Qy       1901 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGC 1960
Db       1872 ACTTCAGCAAGCTCTTCAGCATGGAGAACTGGGGAGGAGCCACGTTTGACGTCGCCATGC 1931
Qy       1961 GTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 2020
Db       1932 GCTTCCTGTATGAGTGCCCCTGGCGGCGGCTGCAGGAGCTCCGGGAGCTCATCCCCAACA 1991
Qy       2021 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080
Db       1992 TCCGTTTCCAGATGCTGCTGCGGGGGGCCAATGCTGTGGGCTACACCAACTACCCAGACA 2051
Qy       2081 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 2140
Db       2052 ACGTGGTCTTCAAGTTCTGTGAAGTGGCCAAAGAGAATGGCATGGATGTCTTCCGTGTGT 2111
Qy       2141 TCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCA 2200
Db       2112 TTGACTCCCTCAACTACTTGCCCAACATGCTGCTGGGCATGGAGGCGG---CAGGAAGTG 2168
Qy       2201 ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAA 2260
Db       2169 CCGGAGGCGTGGTGGAGGCTGCCATCTCATACACGGGCGACGTGGCCGACCCCAGCCGCA 2228
Qy       2261 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2320
Db       2229 CCAAGTACTCACTGCAGTACTACATGGGCTTGGCTGAAGCGCTGGTGCGAGCTGGCACCC 2288
Qy       2321 ACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2380
Db       2289 ACATCCTGTGCATCAAGGACATGGCCGGGCTGCTGAAGCCCACGGCCTGCACCATGCTGG 2348
Qy       2381 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTG 2437

Db       2349 TCAGCTCCCTCCGGGACCGCTTCCCCGACCTCCCACTGCACATCCACACCCACGCACCGT 2408
Qy       2438 CGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497
Db       2409 CAGGGGCAGGCGTGGCAGCCATGCTGGCCTGTGCCCAGGCTGGAGCTGATGTGGTGGATG 2468
Qy       2498 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTG 2557
Db       2469 TGGCAGCTGATTCCATGTCTGGGATGACTTCACAGCCCAGCATGGGGGCCCTGGTGGCCT 2528
Qy       2558 CATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGC 2617
Db       2529 GTACCAGAGGGACTCCCCTGGACACAGAGGTGCCCATGGAGCGCGTGTTTGACTACAGTG 2588
Qy       2618 CGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAA 2677
Db       2589 AGTACTGGGAGGGGGCTCGGGGACTGTACGCGGCCTTCGACTGCACGGCCACCATGAAGT 2648
Qy       2678 CCGG------TCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTG 2731
Db       2649 CTGGCAACTCGGACGTGTATGAAAATGAGATCCCAGGGGGCCAGTACACCAACCTGCACT 2708
Qy       2732 CACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAG 2791
Db       2709 TCCAGGCCCACAGCATGGGGCTTGGCTCCAAGTTCAAGGAGGTCAAGAAGGCCTATGTGG 2768
Qy       2792 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCG 2851
Db       2769 AGGCCAACCAGATGCTGGGCGATCTCATCAAGGTGACGCCCTCCTCCAAGATCGTGGGGG 2828
Qy       2852 ACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCAC 2911
Db       2829 ACCTGGCCCAGTTTATGGTGCAGAATGGATTGAGCCGGGCAGAGGCCGAAGCTCAGGCGG 2888
Qy       2912 AAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTC 2971
Db       2889 AAGAGCTGTCCTTTCCCCGCTCCGTGGTGGAGTTCCTGCAGGGCTACATCGGTGTCCCCC 2948
Qy       2972 CAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCC 3016
Db       2949 ATGGGGGTTTCCCCGAACCCTTTCGCTCTAAGGTACTGAAGGACC 2993
```

RESULT 12
US-10-369-493-31453
; Sequence 31453, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31453
; LENGTH: 3432
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-10-369-493-31453

Query Match 18.2%; Score 659; DB 15; Length 3432;
Best Local Similarity 51.8%; Pred. No. 1.6e-190;
Matches 1782; Conservative 0; Mismatches 1580; Indels 75; Gaps 10;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
```

```
Db        1 AAGATCCTCATCGCCAACCGTGGTGAGATCGCCATCCGGGTGATGCGGGCGGCAAACGAG 60
Qy      298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db       61 ATGGGCAAGAAGACGGTCGCCGTCTATGCCGAGGAAGACAAGCTCTCGCTCCACCGCTTC 120
Qy      358 TTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGAC 417
Db      121 AAGGCGGACGAGGCCTACCGGATCGGCGAGGGCCTCTCGCCCGTGGGCGCCTATCTCTCG 180
Qy      418 ATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGGGATAC 477
Db      181 ATCCCCGAGATCATCCGCGTGGCGCAGATGTCGGGCGCCGATGCGATCCACCCGGGCTAC 240
Qy      478 GGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACTTTT 537
Db      241 GGGCTTCTGTCGGAGAACCCCGATTTCGTCGAGGCCTGCGACGCAGCGGGCATCGCCTTC 300
Qy      538 ATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCC 597
Db      301 ATCGGCCCGAAGGCCGAGACGATGCGCGCGCTCGGCGACAAGGCTTCGGCCCGCCGCGTG 360
Qy      598 GCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCC---ACCCCGAGCAAAAACATCGAT 654
Db      361 GCCATGGCCGCGGGCGTGCCGGTGATCCCGGCGACCGAGGTGCTGGGCGACGATATGGAG 420
Qy      655 GAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGT 714
Db      421 GAGATCAAGCGGCAGGCGGCCGAGATCGGCTATCCGCTGATGCTCAAGGCCTCGTGGGGC 480
Qy      715 GGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACA 774
Db      481 GGCGGTGGGCGCGGCATGCGGCCCATCACTTCCGAGGCCGAGCTGGCCGACAAGGTGCGC 540
Qy      775 GAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCT 834
Db      541 GAGGGCCGACGCGAGGCCGAGGCCGCCTTCGGCAACGGCGAGGGCTATCTCGAGAAGATG 600
Qy      835 GTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTA 894
Db      601 ATCCAGCGCGCGCGCCACGTCGAGGTGCAGATCCTCGGCGACAAATACGGCGCGATCTAC 660
Qy      895 CACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCG 954
Db      661 CATCTCTACGAGCGCGACTGCACCGTGCAGCGGCGCAACCAGAAGGTCGTCGAGCGCGCG 720
Qy      955 CCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTC 1014
Db      721 CCCGCGCCTTATCTCACCGAGGAGCAGCGGACAGAGATCTGCGAGCTCGGCCGCCGGATC 780
Qy     1015 TGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGAT---GAAAAG 1071
Db      781 TGCGCCCATGTGAATTACGAATGCGCGGGCACGGTCGAGTTCCTGATGGATATGGACTCG 840
Qy     1072 GGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAA 1131
Db      841 GAGAAATTCTATTTCATCGAGGTCAATCCGCGGGTTCAGGTCGAACATACCGTCACCGAG 900
Qy     1132 GAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTG 1191
Db      901 GAAGTGACGGGCATCGACATCGTGCAATCCCAGATCCGCATCGCGGAAGGGGCGACGCTG 960
Qy     1192 AAGGAATTGG------GTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAG 1245
Db      961 GCCGAGGCAACCGGATGCCCGAGCCAGGACGACATCAAGCTCTCGGGCCATGCGCTGCAG 1020
Qy     1246 TGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACC 1305
Db     1021 TGCCGCGTCACGACCGAGGATCCGCAGAACAATTTCATTCCCGACTACGGCCGGCTCACC 1080
Qy     1306 GCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGT 1362
```

```
Db     1081 GCCTATCGCTCGGCCACGGGCATGGGCATCCGGCTGGACGGCGGCACGGCCTATGCCGGG 1140
Qy     1363 GGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGAC 1422
Db     1141 GGCGTCATCACCCGCTATTACGATTCACTCCTCGTGAAGGTGACGGCCTGGGCCCCCACG 1200
Qy     1423 TTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTT 1482
Db     1201 CCCGAAAAGGCCATCGCGCGGATGGACCGGGCGCTGCGCGAGTTCCGCATCCGGGGCGTG 1260
Qy     1483 GCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGC 1542
Db     1261 GCCACCAACATCGCCTTCGTCGAGAACCTGCTGAAGCACCCGAGTTTCCTCGACTATTCC 1320
Qy     1543 ATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGAT 1602
Db     1321 TACACGACGAAATTCATCGACACGACGCCCGACCTCTTCAACTTCAAGCCGCGTCGTGAC 1380
Qy     1603 GAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTC------AT 1656
Db     1381 CGGGCGACGAAGATCCTGACCTACATCGCCGACATCACCGTGAACGGGCACCCCGAGACG 1440
Qy     1657 GGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTG 1716
Db     1441 GCCGGCCGGGTCCGCCCCTCGGCCGAGCTGAAGGATCCGAAGGCACCCGAGCCCAAGGGC 1500
Qy     1717 CCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGT 1776
Db     1501 GCACCGCAGCCCGGCACGCGGACCCTCCTCGAGGAGAAGGGCCCGCAGGCGGTGGCCGAC 1560
Qy     1777 GATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAG 1836
Db     1561 TGGATGGCCGCGCAGACCCGCGTGCTGATGACCGACACGACGATGCGCGACGGCCACCAG 1620
Qy     1837 TCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCA 1896
Db     1621 AGCCTGCTCGCCACCCGGATGCGCTCGATCGACATGATCAAGGTCACGCCCGCCTATGCG 1680
Qy     1897 AAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCG 1956
Db     1681 GCGAACCTCGGCGGCCTCTTCTCGGTGGAATGCTGGGGCGGCGCCACCTTCGATGTGGCC 1740
Qy     1957 ATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCG 2016
Db     1741 TATCGCTTCCTGCAGGAATGTCCCTGGCAGCGGCTGCGCGACATCCGCGCCCGGCTGCCC 1800
Qy     2017 AATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCA 2076
Db     1801 AACGTCATGACGCAGATGCTGCTGCGCGCCTCGAACGGGGTCGGCTATACGAACTACCCC 1860
Qy     2077 GACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGC 2136
Db     1861 GACAATGTGGTGCAGGAATTCGTCCGTCAGGCGGCCGAAACCGGCGTCGACGTGTTCCGC 1920
Qy     2137 ATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAG 2196
Db     1921 GTGTTCGACTCGCTCAACTGGGTCGAGAACATGCGCGTGGCAATGGATGCGGTGATCGAG 1980
Qy     2197 ACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAAT 2256
Db     1981 GCCAACA---AGGTCTGCGAGGGCACGATCTGCTACACGGGCGACCTGCTCGATCCCGAC 2037
Qy     2257 GAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGC 2316
Db     2038 CGCTCCAAATACGATCTGAACTACTATGTCGGCATGGGCCGCGCGCTGCGCGACGCGGGC 2097
Qy     2317 GCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAG 2376
Db     2098 GCGCATGTGCTGGGTCTGAAGGACATGGCCGGGCTTCTGAAGCCTGCCGCCGCCCGCGTG 2157
Qy     2377 CTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACT 2436
Db     2158 CTGGTGAAGGCGCTGAAGGAGGAGGTGGGCCTGCCGATCCACTTCCACACCCACGACACG 2217
```

```
Qy       2437 GCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGAC 2496
Db       2218 AGCGGAATTGCGGGCGCGACCGTGCTCGCCGCCTGCGATGCCGGGGTGGATGCGGTCGAT 2277

Qy       2497 GGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCT 2556
Db       2278 GCGGCGATGGACGCCTTCTCGGGCGGCACCTCGCAGCCCTGCCTCGGCTCGATCGTCGAG 2337

Qy       2557 GCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAG 2616
Db       2338 GCGCTGAAGCACACCGACCGCGACACGGGCCTCGACATCGCGGCGATCCGCGAGATCTCG 2397

Qy       2617 CCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCA 2676
Db       2398 GACTACTGGGGCCATGTGCGCCAGCAGTATTCGGCCTTCGAATCGGGCCTGCCCTCGCCC 2457

Qy       2677 ACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAG 2736
Db       2458 GCCTCCGAGGTCTATCTGCACGAGATGCCCGGCGGCCAGTTCACCAACCTCAAGGCGCAG 2517

Qy       2737 GCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTT 2796
Db       2518 GCGCGCTCCATGGGGCTCGAGGAACGCTGGTCCGAGGTGGCGCAGGCCTATGCCGACGCG 2577

Qy       2797 AATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCGACCTC 2856
Db       2578 AACCGGATGTTCGGCGACATCGTGAAGGTCACGCCCTCGTCGAAGGTGGTGGGCGACATG 2637

Qy       2857 GCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAG 2916
Db       2638 GCGCTGATGATGGTGGCGCAGGGGCTGACGCGGGAAGAGGTCGAGGATCCCGAGGTCGAG 2697

Qy       2917 TACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGT 2976
Db       2698 GTGAGCTTCCCGGACTCGGTCGTGGACATGCTCAAGGGCAATCTCGGCCAGCCCCACGGC 2757

Qy       2977 GGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTC---------------- 3020
Db       2758 GGCTGGCCCGAGCCGATCCTGAAGAAGGTGCTGAAGGGCGAGGCACCCTCGACCGAGCGA 2817

Qy       3021 -------------------------CGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAG 3054
Db       2818 CCGGGCGCGCATCTGCCGCCCGTCGACATCGCAGCCGCGCGCGAGAAGCTCCTGTCGGAG 2877

Qy       3055 GAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCG- 3113
Db       2878 ATCAAGCAGGGCGACGACGATCCGCTCGACACGGCCGTCGATGCCGAGGATCTGAACGGC 2937

Qy       3114 --CCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAAC 3171
Db       2938 TACCTCATGTATCCCAAGGTCTTCACCGACTACCGCGCCCGCCACAGGATCTACGGGCCG 2997

Qy       3172 ACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTG 3231
Db       2998 GTGCGGACGCTGCCGACCCGGACCTTCTTCTACGGGATGGAGCCGGGCGAGGAGATCTCG 3057

Qy       3232 ATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCA 3291
Db       3058 GCCGAAAT---CGACCCCGGCAAGACGCTCGAGATCCGGCTGTCCGCCGTGGGCGAGACG 3114

Qy       3292 GACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGT 3351
Db       3115 TCGGACGACGGCGATGCCAAGGTGTTCTTCGAGCTGAACGGCCAGCCGCGGGTGATCCGC 3174

Qy       3352 GTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAAC 3411
Db       3175 GTGGCCAACCGGGCGGTGAAGGCCAAGACCGCGACCCGGCCCAAGGCGCAGGACGGCAAT 3234

Qy       3412 AAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACC---GTGACTGTTGCTGAAGGT 3468
Db       3235 CCGGCCCATGTCGGCGCCCCGATGCCGGGTTCGGTCGCCTCGGTCGCGGTCTCGGCGGGC 3294

Qy       3469 GATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3528
Db       3295 CAGAAGGTGAAGCCGGGCGATCTCCTCGTCACCATCGAGGCGATGAAGATGGAGACCGGG 3354

Qy       3529 ATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTG 3588
Db       3355 CTCCATGCCGACCGTGCGGCCACCGTGAAGGCTGTCCATGTCGGGCCCGGCGCGCAGATC 3414

Qy       3589 GAAGGTGGCGACTTGAT 3605
Db       3415 GAGGCGAAGGACCTGCT 3431
```

RESULT 13
US-10-369-493-35714
; Sequence 35714, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35714
; LENGTH: 3459
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35714

Query Match 17.8%; Score 643.6; DB 15; Length 3459;
Best Local Similarity 51.6%; Pred. No. 8.7e-186;
Matches 1788; Conservative 0; Mismatches 1599; Indels 75; Gaps 11;

```
Qy        229 GCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCA 288
Db          4 GCCATCACGAAGATCCTCGTCGCCAACCGGTCAGAAATCGCCATCCGCGTCTTTCGCGCG 63

Qy        289 GCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTC 348
Db         64 GCCAATGAACTGGGCCTCAAAACCGTGGCGATCTGGGCCGAGGAAGACAAATATTCGCTG 123

Qy        349 CACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACC-----------------GAA 390
Db        124 CACCGCTTCAAGGCCGACGAGAGCTACCAGGTCGGGCGCGGTCCGCATCTCACCAAGGAT 183

Qy        391 GGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTT 450
Db        184 ATGGGGCCGATCGAGAGCTATCTGTCGATCGAGGAGGTGATCCGCGTCGCCCGGCTCTCG 243

Qy        451 AAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGC 510
Db        244 GGCGCCGATGCCATCCACCCCGGCTACGGGCTTCTGTCCGAAAGCCCCGAATTCGCCGAA 303

Qy        511 GAGTGTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACC 570
Db        304 GCCTGCGCGCAAGCGGGCATCACCTTCATCGGGCCGAAGCCGGACACGATGCGCCGCCTC 363

Qy        571 GGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAA 630
Db        364 GGCAACAAGGTCGCGGCGCGCAATCTCGCCATCGAGGTCGGCGTGCCTGTCATTCCCGCC 423

Qy        631 TCCACCCC---GAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687
Db        424 ACCGATCCGTTGCCGGACGATATGGAGGCGGTCAAGAAACTGGCCAAGGAGATCGGCTAT 483
```

```
Qy    688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCA 747
Db    484 CCAGTGATGCTCAAGGCCTCCTGGGGCGGCGGTGGGCGCGGCATGCGCGCCATCCGCTCC 543

Qy    748 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807
Db    544 GAGGCCGATCTCGCCCGTGAGGTCACGGAAGGCAAGCGCGAGGCGAAAGCCGCCTTCGGC 603

Qy    808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867
Db    604 AAGGACGAGGTCTATCTCGAAAAGCTGATCGAGCGCGCCCGCCACGTCGAGGTGCAGGTG 663

Qy    868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927
Db    664 CTTGGCGACACGCATGGCAATGCCGTGCACCTGTTCGAGCGCGACTGCTCGATCCAGCGC 723

Qy    928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987
Db    724 CGCAACCAGAAGGTCGTCGAGCGGGCGCCGGCGCCCTATCTCGAAATGTCGCAGCGCGAG 783

Qy    988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047
Db    784 GAGCTTTGCGGCTACGCGCTGAAGATCGCGCGCGAGACCAGCTATATCGGCGCCGGCACG 843

Qy   1048 GTGGAATTCTTGGTCGAT---GAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGT 1104
Db    844 GTCGAGTTCCTGCAGGATGCCGATACCGGAAAATTCTATTTCATCGAGGTCAACCCGCGC 903

Qy   1105 ATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAG 1164
Db    904 ATCCAGGTCGAGCATACCGTCACCGAGCAGGTGACCGGCATCGATATCGTCAAGGCGCAG 963

Qy   1165 ATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG---------GGTCTGACCCAAGAT 1215
Db    964 ATCCACATCCTCGACGGCTTCGCCATCGGCACACCGCAATCGGGCGTGCCGGCGCAGAAG 1023

Qy   1216 AAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAAC 1275
Db   1024 GATATCAGGCTGAACGGCCATGCCTTGCAGTGCCGCATCACCACCGAGGATCCCGAGCAC 1083

Qy   1276 GGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTT 1335
Db   1084 AATTTCATCCCGGACTATGGCCGCATCACCGCCTATCGCGGCGCCACCGGCTTCGGCATC 1143

Qy   1336 CGTCT---TGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATG 1392
Db   1144 CGGCTGGATGGCGGCACCGCCTATTCTGGCGCGGTCATCACCCGCTTCTACGATCCGCTG 1203

Qy   1393 CTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGC 1452
Db   1204 CTGGAGAAGGTGACGGCGTGGGCGCCGACGCCGGCCGAGACCATTGCCCGCATGAACCGG 1263

Qy   1453 GCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTG 1512
Db   1264 GCGCTGCGCGAATTCCGCATTCGCGGCGTTGCCACCAACCTCACCTTCCTCGAAGCGATC 1323

Qy   1513 CTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCG 1572
Db   1324 ATCAACCACCCGAGCTTCGCCGACAATTCCTACACGACCAAGTTCATCGACACGACGCCG 1383

Qy   1573 CACCTCCTTCAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCA 1632
Db   1384 GAGCTGTTCCAGCAGGTCAAGCGGCAGGACCGCGCGACCAAGCTGCTCAACTATCTGGCC 1443

Qy   1633 GATGTCACCGTGAACAAGCCTCATG------GTGTGCGTCCAAAGGATGTTGCAGCTCCT 1686
Db   1444 GATGTCAGCGTCAACGGCCATCCCGAGACGCGCGGCCGGCCGACGCCGAAGGCCGATGCG 1503

Qy   1687 ATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTG 1746
Db   1504 GCCGCACCGGTCGTGCCCTATCTCAACGGCAATGTGCCGGGCGGCAGCAAGCAGAAGCTG 1563

Qy   1747 AAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTT 1806
Db   1564 GACGTGCTCGGTCCGGAAAAATTCGCCGGCTGGATGCGCGAACAAAGGCAAGTTCTGGTC 1623

Qy   1807 ACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTC 1866
Db   1624 ACCGATACGACGATGCGCGACGGCCACCAGTCGCTGCTCGCGACGCGCATGCGCACGCAC 1683

Qy   1867 GCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAG 1926
Db   1684 GACATTGCCAATATTGCCGGCACCTATGCGCGCGCCCTGCCGCAGCTTCTGTCGCTCGAA 1743

Qy   1927 GCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGAC 1986
Db   1744 TGCTGGGGCGGCGCGACCTTCGACGTCGCCATGCGCTTCCTTACCGAGGATCCTTGGGAG 1803

Qy   1987 AGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGC 2046
Db   1804 CGGCTGTCGCTGGTGCGCGAGGCGGCGCCCAACCTGTTGCTGCAGATGCTGCTGCGCGGC 1863

Qy   2047 CGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAA 2106
Db   1864 GCCAACGGCGTCGGTTACACCAATTATCCCGACAATGTCGTGCAGCATTTCGTCAAGCAG 1923

Qy   2107 GCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAG 2166
Db   1924 GCGGCGAGCGGCGGTATCGACCTGTTCCGCGTCTTCGACTGCCTGAACTGGGTCGAGAAC 1983

Qy   2167 ATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATG 2226
Db   1984 ATGCGCGTCGCCATGGACGCGGTC---GGCGCCGAAGGCAAGCTGATCGAAGCGGCGATG 2040

Qy   2227 GCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTA 2286
Db   2041 TGCTACACCGGCGACATTCTCGATCCGGCGCGGGCCAAATACGATCTGAAATATTATGTC 2100

Qy   2287 AAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCT 2346
Db   2101 GGCCTGGCCAGCGAATTGCAGGCCGCCGGCGCCCACATCATCGCCGTCAAGGACATGGCC 2160

Qy   2347 GGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGAT 2406
Db   2161 GGCCTGTTGAAGCCCGCCGCCGCCCGCGTGCTGTTCAAGGCGCTGCGCGAGGCGACCGAC 2220

Qy   2407 CTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCT 2466
Db   2221 CTGCCGATCCATTTCCATACCCACGACACGTCTGGCCTGTCGGCGGCGACCGTGCTGGCG 2280

Qy   2467 GCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACC 2526
Db   2281 GCGGTTGAGAGCGGCGTCGACGCCATCGATGCGGCGATGGACGCCTTCTCCGGCAACACG 2340

Qy   2527 TCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGT 2586
Db   2341 TCGCAGCCTTGCCTGGGCTCGATCGTCGAGGCGCTGAAGGGCACCGAGCGCGACCCGGGC 2400

Qy   2587 TTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTAC 2646
Db   2401 CTCGACCCGCAATGGATCCGCAAGATCTCGTTCTACTGGGAAGCGGTGCGCAACCAGTAC 2460

Qy   2647 CTGCCATTTGAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCA 2706
Db   2461 GCCGCCTTCGAAAGCGACCTCAAAGGGCCGGCTTCGGAAGTCTACCTGCATGAAATGCCG 2520

Qy   2707 GGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTC 2766
Db   2521 GGCGGACAATTCACCAACCTCAAGGAGCAGGCGCGCTCGCTCGGGCTGGAAACGCGCTGG 2580

Qy   2767 GAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTC 2826
Db   2581 CATGAGGTGGCGCAGACCTATCACGACGTCAATCTGATGTTCGGCGACATCGTCAAGGTG 2640

Qy   2827 ACCCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGAT 2886
```

```
Db       2641 ACGCCGTCATCCAAGGTCGTCGGCGACATGGCCCTGATGATGGTCAGCCAGGACCTGACC 2700
Qy       2887 CCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTC 2946
Db       2701 GTCGCCGATGTCGAGAACCCGGCCAGGGATATTGCCTTCCCGGACTCGGTCGTCTCGATG 2760
Qy       2947 CTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCA 3006
Db       2761 CTGCGCGGCGATCTCGGCCAGTCGCCCGGCGGTTGGCCGGCGGCGCTGCAGAAGAAGGCG 2820
Qy       3007 CTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCG 3066
Db       2821 CTGAAAGGCGACGAGCCGATCACGGCACGGCCAGGCTCGCTGCTCAAGCCCGCCGATCTC 2880
Qy       3067 CACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAG----------C 3105
Db       2881 AAGGCCAGCCGCAAGGAGATCGAGGAGAAGCTCGAGCGCAAGCTCTCGGAATACGAATTC 2940
Qy       3106 CTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTC 3165
Db       2941 GCCTCCTGGCTGATGTATCCGAAGGTGTTCACTGACTTCGCCGGCGCGCAGGAAACCTAC 3000
Qy       3166 GGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAG 3225
Db       3001 GGCCCGGTCAGCGTCCTGCCGACGCCGACCTATTTCTACGGCATGAAATCGGAAGACGAG 3060
Qy       3226 ACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCT 3285
Db       3061 ATCTTCATCGACAT---CGAGAAGGGTAAGACGCTGGTCGTGCGGTGCCTTGCCATCGGC 3117
Qy       3286 GAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCA 3345
Db       3118 GATGTGGACGACAAGGGCATGGTCACGGTGTTCTTCGAGCTCAACGGCCAGCCACGGCGC 3177
Qy       3346 ATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACCGCAGAA---AAGGCAGAT 3402
Db       3178 GTGAAGGTGCCGGACCGGGCGCATGGCGCCTCCGCCGCCAAGGCGCGCCGTAAGGCCGAG 3237
Qy       3403 TCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACT---GTT 3459
Db       3238 CCTGGCAACGAGGCGCATGTCGGCGCGCCGATGCCGGGCGTGGTTTCGGCGCTTTCGGTC 3297
Qy       3460 GCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATG 3519
Db       3298 GCCACGGGGCAAGCGGTGAAAGCCGGTGACGTGCTGCTGTCCATCGAGGCGATGAAGATG 3357
Qy       3520 GAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCA 3579
Db       3358 GAAACGGCGCTGCATGCCGAGCGTGACGGCACCGTCGCCGAGGTGCTGGTCAAGGCCGGC 3417
Qy       3580 ACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAA 3621
Db       3418 GACCAGATCGATGCCAAGGATCTGCTGATAGCCTTCGCTTGA 3459
```

RESULT 14
US-10-369-493-35137
; Sequence 35137, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 35137
; LENGTH: 3441
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35137

Query Match 17.0%; Score 617; DB 15; Length 3441;
Best Local Similarity 51.6%; Pred. No. 1.2e-177;
Matches 1776; Conservative 0; Mismatches 1590; Indels 75; Gaps 13;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db          7 AAAATACTTGTTGCCAACCGATCCGAAATTGCGATCCGCGTTTTCCGGGCAGCGAACGAG 66
Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db         67 CTTGGGATCAAAACCGTTGCGATTTGGGCGGAAGAGGACAAGCTGTCTTTGCACCGCTTC 126
Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACC----------------GAAGGCTCACCA 399
Db        127 AAGGCGGATGAATCCTATCAGGTCGGCAGGGGTCCGCATCTCGCCAAGGATATGGGACCG 186
Qy        400 GTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGAT 459
Db        187 ATCGAGAGTTATCTTTCGATCGAGGAGGTTATCCGCGTGGCCAAGCTCTCCGGAGCGGAT 246
Qy        460 GCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCG 519
Db        247 GCGATCCATCCCGGTTATGGTCTTCTGTCCGAAAGCCCTGAGTTCGTGGAGGCCTGCAAC 306
Qy        520 GAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAG 579
Db        307 AAGGCCGGCATCACCTTCATCGGGCCGACACCGGACACGATGCGCCAGCTCGGCAACAAG 366
Qy        580 TCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCA---CC 636
Db        367 GTCGCCGCCCGCAATCTGGCGATTTCGGTTGATGTCCCCGTTGTTCCCGCCACCAATCCT 426
Qy        637 CCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTT 696
Db        427 CTGCCTGACGATATCGCGGAAGTGGAGCGTATGGCCGAGGAAATCGGTTATCCCGTCATG 486
Qy        697 GTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAG 756
Db        487 CTCAAGGCCTCCTGGGGCGGCGGTGGCCGCGGCATGCGCGCCATTCGCAAAAAGGAAGAT 546
Qy        757 CTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCG 816
Db        547 CTCGCCCGTGAGGTGACGGAAGCCAAGCGCGAGGCGAAGGCCGCCTTCGGCAAGGACGAG 606
Qy        817 GTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGAT 876
Db        607 GTTTATCTGGAAAAGCTCGTCGAGCGCGCCCGCCATGTCGAGAGCCAGATTCTCGGCGAT 666
Qy        877 CACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAA 936
Db        667 ACGCATGGCAATGTCGTGCATCTGTTCGAGCGGGATTGTTCGATCCAGAGACGCAATCAA 726
Qy        937 AAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGT 996
Db        727 AAGGTCGTCGAGCGCGCGCCAGCACCGTATCTGTCCGAAGCGCAGCGCCAGGAACTGGCC 786
Qy        997 GCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTC 1056
Db        787 GCCTATTCCCTGAAGATCGCAGCCGCGACCAACTACATCGGCGCCGGCACTGTCGAATAT 846
Qy       1057 TTGGTCGAT---GAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTT 1113
Db        847 CTGATGGATGCCGATACCGGCAAATTCTACTTCATCGAGGTCAATCCGCGTATCCAGGTG 906
Qy       1114 GAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTG 1173
Db        907 GAGCATACCGTTACCGAAGTCGTGACCGGCATCGATATCGTCAAGGCGCAGATCCATATT 966
```

```
Qy      1174 GCTGCTGGTGCAACCTT------GAAGGAATTGGGTCTGACC---CAAGATAAGATCAAG 1224
Db       967 CTCGAAGGAGCGGCCATCGGCACGGCGGAATCCGGCGTGCCCAAGCAGGAAGATATCCGC 1026

Qy      1225 ACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGC 1284
Db      1027 CTCAACGGCCATGCGCTGCAGTGCCGCATCACGACGGAAGACCCGGAACACAATTTCATT 1086

Qy      1285 CCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGAC 1344
Db      1087 CCGGACTATGGCCGTATTACCGCCTATCGCTCGGCTTCCGGTTTCGGCATCCGTCTGGAT 1146

Qy      1345 GGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAA 1401
Db      1147 GGGGGCACGTCCTATACCGGCGCCGTCATCACCCGTTATTACGATCCGCTGCTCGTCAAG 1206

Qy      1402 ATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCT 1461
Db      1207 GTCACGGCCTGGGCGCCGGAGCCGGATGAAGCGATCAGCCGCATGGATCGCGCGTTGCGC 1266

Qy      1462 GAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAA 1521
Db      1267 GAATTTCGTATCCGCGGCGTTGCGACCAACCTGACCTTCCTCGAAGCCATCATCGGTCAC 1326

Qy      1522 GAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTT 1581
Db      1327 GACAGTTTCCGCAACAACACCTATACGACGCGCTTCATCGATTCGACGCCGGAGTTGTTC 1386

Qy      1582 CAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACC 1641
Db      1387 GCGCAGGTCAAGCGTCAGGACCGCGCCACCAAGCTTCTGACCTATCTGGCCGATGTGACC 1446

Qy      1642 GTGAACAAGCCTCATGGT------GTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAG 1695
Db      1447 GTCAATGGTCACCCAGAAACCAAGGGTCGTGCGAAGCCTGCGGACAAGGCGGCAAAGCCC 1506

Qy      1696 CTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTT 1755
Db      1507 ATCGTGCCCTATATCGATGCGCCGACGCCTGACGGCACCAAGCAATTGCTGGACAAGCTT 1566

Qy      1756 GGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACC 1815
Db      1567 GGCCCTAAGGGCTTTGCGGACTGGATGCGCAACGAAAAGCGCGTGCTGGTCACGGATACG 1626

Qy      1816 ACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAG 1875
Db      1627 ACCATGCGCGACGGACACCAGTCGCTGCTGGCCACCCGCGTCCGTACCCATGATATCGCC 1686

Qy      1876 CCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGC 1935
Db      1687 CGCGTTGCCAGCGTTTATTCCAAGGCGCTTCCGCAGCTTCTGTCGCTGGAATGCTGGGGC 1746

Qy      1936 GGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGAC 1995
Db      1747 GGTGCGACCTTCGACGTCTCAATGCGCTTCCTGACCGAAGACCCGTGGGAGCGCCTGTCA 1806

Qy      1996 GAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACC 2055
Db      1807 CTCATTCGCGAAGGTGCGCCGAACCTGCTGTTGCAGATGCTTCTGCGTGGTGCCAACGGC 1866

Qy      2056 GTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGC 2115
Db      1867 GTCGGTTACAAGAACTATCCCGATAATGTCGTGAAGTATTTCGTCCGTCAGGCGGCAAGG 1926

Qy      2116 TCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCA 2175
Db      1927 GGCGGCGTCGATCTTTTTCGCGTCTTCGACTGCCTGAACTGGGTGGAGAATATGCGGGTG 1986

Qy      2176 GCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCT 2235
Db      1987 TCGATGGATGCGATTGCCGAGGAGAACA---AGCTCTGCGAGGCAACCATCTGCTACACC 2043
```

```
Qy      2236 GGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCA 2295
Db      2044 GGCGATCTCTTGAATTCGGCACGTCCGAAATATGATCTCAAATATTATACCAACCTCGCG 2103

Qy      2296 GAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTT 2355
Db      2104 GTCGAGCTGGAAAAGGCCGGCGCCCATATCATCGCCGTCAAGGATATGGCGGGCCTGCTG 2163

Qy      2356 CGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTG 2415
Db      2164 AAACCGGCCGCCGCCAAGGTTCTGTTCAAGGCGTTGCGCGAGGCGACCGGTCTGCCGATC 2223

Qy      2416 CACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAA 2475
Db      2224 CACTTCCACACCCATGATACGTCGGGCATTTCCGCCGCGACCGTTCTTGCGGCCGTGGAT 2283

Qy      2476 GCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCA 2535
Db      2284 GCCGGCGTCGACGCCGTCGATGCCGCGATGGATGCCTTCTCCGGCAATACGTCGCAGCCC 2343

Qy      2536 TCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTC 2595
Db      2344 TGCCTCGGCTCGATCGTGGAGGCGCTCTCCGGATCGGAACGCGATACGGGTCTCGATACC 2403

Qy      2596 GAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTT 2655
Db      2404 GAATGGATTCGCCGCATTTCCTTCTATTGGGAAGCCGTGCGCAATCAATATGCGGCGTTC 2463

Qy      2656 GAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAG 2715
Db      2464 GAGAGTGACCTCAAAGGGCCTGCCTCGGAAGTCTATCTGCATGAAATGCCGGGTGGCCAG 2523

Qy      2716 TTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATC 2775
Db      2524 TTCACCAACCTCAAGGAACAGGCACGTTCGCTGGGTCTCGAAAGCCGCTGGCACGAGGTG 2583

Qy      2776 GAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCC 2835
Db      2584 GCGCAGGCCTATGCCGATGCCAACAGGATGTTCGGCGATATCGTCAAGGTGACGCCGTCC 2643

Qy      2836 TCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGAC 2895
Db      2644 TCCAAGGTAGTCGGCGACATGGCGCTGATGATGGTGAGCCAGGACCTGACGGTGGCCGAT 2703

Qy      2896 TTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGC 2955
Db      2704 GTCGAAAACCCTGACCGTGAAGTGTCCTTCCCGGACTCGGTGGTATCGATGCTGAAGGGC 2763

Qy      2956 GAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGC 3015
Db      2764 GATCTCGGTCAATCGCCGGGCGGCTGGCCGGAAGCGCTGCAGAAGAAGGCGCTGAAGGGT 2823

Qy      3016 CGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGA----------------AG 3058
Db      2824 GAAAAGCCTTACACGGTTCGTCCGGGTTCGCTTCTGGAAGATGCCGATCTCGATGCCGAA 2883

Qy      3059 AGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGC---- 3114
Db      2884 CGGAAGGTGATCGAGACCAAGCTGGAGCGCAAGGTCGATGACTTCGAGTTTGCCTCCTAT 2943

Qy      3115 CTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACC 3174
Db      2944 CTGATGTATCCGAAGGTGTTCACCGATTTCGCGCTGACCGCCGAGACATACGGCCCGGTC 3003

Qy      3175 TCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATC 3234
Db      3004 TCCGTGCTGCCCACCCATGCCTATTTCTATGGCATGGAGGATGGCGAAGAG---CTGTTT 3060

Qy      3235 CGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGAC 3294
Db      3061 GCGGATATCGAGCGCGGCAAGACGCTTGTCATCGTCAATCAGGCTTCGTCCGGCATTGAC 3120

Qy      3295 GATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTG 3354
```

```
Db       3121 GATAAGGGCATGGTGACGGTGTTCTTCGAGATCAACGGCCAGCCGCGCCGCATCAAGGTG 3180
Qy       3355 CGTGACCGCTCC---GTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAAC 3411
Db       3181 CCGGATCGCGCCCATGGCGCTTCCGGCTCTGCGGTGCGCCGCAAGGCAGAACCCGGCAAT 3240
Qy       3412 AAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCAC---CGTGACTGTTGCTGAAGGT 3468
Db       3241 GCGTCCCATATCGGCGCGCCGATGCCTGGTGTCATCAGCCGCGTCTTCATCAACCAGGGT 3300
Qy       3469 GATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3528
Db       3301 CAGGAGGTCAAGGCCGGCGACGTGCTGCTCTCGATCGAGGCCATGAAGATGGAAACGGCG 3360
Qy       3529 ATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTG 3588
Db       3361 CTGCATGCCGAACGCGACGGCAAGATCGCTGAAGTTCTGGTGAAGCCCGGTGACCAGATC 3420
Qy       3589 GAAGGTGGCGACTTGATCGTC 3609
Db       3421 GACGCCAAGGATTTGTTGATC 3441
```

RESULT 15
US-10-369-493-38504
; Sequence 38504, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38504
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38504

Query Match 17.0%; Score 617; DB 15; Length 3453;
Best Local Similarity 51.6%; Pred. No. 1.2e-177;
Matches 1776; Conservative 0; Mismatches 1590; Indels 75; Gaps 13;

```
Qy        238 AAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAA 297
Db         13 AAAATACTTGTTGCCAACCGATCCGAAATTGCGATCCGCGTTTTCCGGGCAGCGAACGAG 72
Qy        298 ACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCT 357
Db         73 CTTGGGATCAAAACCGTTGCGATTTGGGCGGAAGAGGACAAGCTGTCTTTGCACCGCTTC 132
Qy        358 TTTGCTTCTGAAGCTGTCCGCATTGGTACC----------------GAAGGCTCACCA 399
Db        133 AAGGCGGATGAATCCTATCAGGTCGGCAGGGGTCCGCATCTCGCCAAGGATATGGGACCG 192
Qy        400 GTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGAT 459
Db        193 ATCGAGAGTTATCTTTCGATCGAGGAGGTTATCCGCGTGGCCAAGCTCTCCGGAGCGGAT 252
Qy        460 GCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCG 519
Db        253 GCGATCCATCCCGGTTATGGTCTTCTGTCCGAAAGCCCTGAGTTCGTGGAGGCCTGCAAC 312
Qy        520 GAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAG 579
```

```
Db        313 AAGGCCGGCATCACCTTCATCGGGCCGACACCGGACACGATGCGCCAGCTCGGCAACAAG 372
Qy        580 TCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCA---CC 636
Db        373 GTCGCCGCCCGCAATCTGGCGATTTCGGTTGATGTCCCCGTTGTTCCCGCCACCAATCCT 432
Qy        637 CCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTT 696
Db        433 CTGCCTGACGATATCGCGGAAGTGGAGCGTATGGCCGAGGAAATCGGTTATCCCGTCATG 492
Qy        697 GTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAG 756
Db        493 CTCAAGGCCTCCTGGGGCGGCGGTGGCCGCGGCATGCGCGCCATTCGCAAAAAGGAAGAT 552
Qy        757 CTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCG 816
Db        553 CTCGCCCGTGAGGTGACGGAAGCCAAGCGCGAGGCGAAGGCCGCCTTCGGCAAGGACGAG 612
Qy        817 GTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGAT 876
Db        613 GTTTATCTGGAAAAGCTCGTCGAGCGCGCCCGCCATGTCGAGAGCCAGATTCTCGGCGAT 672
Qy        877 CACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAA 936
Db        673 ACGCATGGCAATGTCGTGCATCTGTTCGAGCGGGATTGTTCGATCCAGAGACGCAATCAA 732
Qy        937 AAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGT 996
Db        733 AAGGTCGTCGAGCGCGCGCCAGCACCGTATCTGTCCGAAGCGCAGCGCCAGGAACTGGCC 792
Qy        997 GCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTC 1056
Db        793 GCCTATTCCCTGAAGATCGCAGCCGCGACCAACTACATCGGCGCCGGCACTGTCGAATAT 852
Qy       1057 TTGGTCGAT---GAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTT 1113
Db        853 CTGATGGATGCCGATACCGGCAAATTCTACTTCATCGAGGTCAATCCGCGTATCCAGGTG 912
Qy       1114 GAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTG 1173
Db        913 GAGCATACCGTTACCGAAGTCGTGACCGGCATCGATATCGTCAAGGCGCAGATCCATATT 972
Qy       1174 GCTGCTGGTGCAACCTT------GAAGGAATTGGGTCTGACC---CAAGATAAGATCAAG 1224
Db        973 CTCGAAGGAGCGGCCATCGGCACGGCGGAATCCGGCGTGCCCAAGCAGGAAGATATCCGC 1032
Qy       1225 ACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGC 1284
Db       1033 CTCAACGGCCATGCGCTGCAGTGCCGCATCACGACGGAAGACCCGGAACACAATTTCATT 1092
Qy       1285 CCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGAC 1344
Db       1093 CCGGACTATGGCCGTATTACCGCCTATCGCTCGGCTTCCGGTTTCGGCATCCGTCTGGAT 1152
Qy       1345 GGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAA 1401
Db       1153 GGGGGCACGTCCTATACCGGCGCCGTCATCACCCGTTATTACGATCCGCTGCTCGTCAAG 1212
Qy       1402 ATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCT 1461
Db       1213 GTCACGGCCTGGGCGCCGGAGCCGGATGAAGCGATCAGCCGCATGGATCGCGCGTTGCGC 1272
Qy       1462 GAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAA 1521
Db       1273 GAATTTCGTATCCGCGGCGTTGCGACCAACCTGACCTTCCTCGAAGCCATCATCGGTCAC 1332
Qy       1522 GAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTT 1581
Db       1333 GACAGTTTCCGCAACAACACCTATACGACGCGCTTCATCGATTCGACGCCGGAGTTGTTC 1392
Qy       1582 CAGGCTCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACC 1641
```

```
Db      1393 GCGCAGGTCAAGCGTCAGGACCGCGCCACCAAGCTTCTGACCTATCTGGCCGATGTGACC 1452
Qy      1642 GTGAACAAGCCTCATGGT------GTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAG 1695
Db      1453 GTCAATGGTCACCCAGAAACCAAGGGTCGTGCGAAGCCTGCGGACAAGGCGGCAAAGCCC 1512
Qy      1696 CTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTT 1755
Db      1513 ATCGTGCCCTATATCGATGCGCCGACGCCTGACGGCACCAAGCAATTGCTGGACAAGCTT 1572
Qy      1756 GGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACC 1815
Db      1573 GGCCCTAAGGGCTTTGCGGACTGGATGCGCAACGAAAAGCGCGTGCTGGTCACGGATACG 1632
Qy      1816 ACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAG 1875
Db      1633 ACCATGCGCGACGGACACCAGTCGCTGCTGGCCACCCGCGTCCGTACCCATGATATCGCC 1692
Qy      1876 CCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGC 1935
Db      1693 CGCGTTGCCAGCGTTTATTCCAAGGCGCTTCCGCAGCTTCTGTCGCTGGAATGCTGGGGC 1752
Qy      1936 GGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGAC 1995
Db      1753 GGTGCGACCTTCGACGTCTCAATGCGCTTCCTGACCGAAGACCCGTGGGAGCGCCTGTCA 1812
Qy      1996 GAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACC 2055
Db      1813 CTCATTCGCGAAGGTGCGCCGAACCTGCTGTTGCAGATGCTTCTGCGTGGTGCCAACGGC 1872
Qy      2056 GTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGC 2115
Db      1873 GTCGGTTACAAGAACTATCCCGATAATGTCGTGAAGTATTTCGTCCGTCAGGCGGCAAGG 1932
Qy      2116 TCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCA 2175
Db      1933 GGCGGCGTCGATCTTTTTCGCGTCTTCGACTGCCTGAACTGGGTGGAGAATATGCGGGTG 1992
Qy      2176 GCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCT 2235
Db      1993 TCGATGGATGCGATTGCCGAGGAGAACA---AGCTCTGCGAGGCAACCATCTGCTACACC 2049
Qy      2236 GGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCA 2295
Db      2050 GGCGATCTCTTGAATTCGGCACGTCCGAAATATGATCTCAAATATTATACCAACCTCGCG 2109
Qy      2296 GAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTT 2355
Db      2110 GTCGAGCTGGAAAAGGCCGGCGCCCATATCATCGCCGTCAAGGATATGGCGGGCCTGCTG 2169
Qy      2356 CGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTG 2415
Db      2170 AAACCGGCCGCCGCCAAGGTTCTGTTCAAGGCGTTGCGCGAGGCGACCGGTCTGCCGATC 2229
Qy      2416 CACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAA 2475
Db      2230 CACTTCCACACCCATGATACGTCGGGCATTTCCGCCGCGACCGTTCTTGCGGCCGTGGAT 2289
Qy      2476 GCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCA 2535
Db      2290 GCCGGCGTCGACGCCGTCGATGCCGCGATGGATGCCTTCTCCGGCAATACGTCGCAGCCC 2349
Qy      2536 TCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTC 2595
Db      2350 TGCCTCGGCTCGATCGTGGAGGCGCTCTCCGGATCGGAACGCGATACGGGTCTCGATACC 2409
Qy      2596 GAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTT 2655
Db      2410 GAATGGATTCGCCGCATTTCCTTCTATTGGGAAGCCGTGCGCAATCAATATGCGGCGTTC 2469
Qy      2656 GAGTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAG 2715
Db      2470 GAGAGTGACCTCAAAGGGCCTGCCTCCGGAAGTCTATCTGCATGAAATGCCGGGTGGCCAG 2529

Qy      2716 TTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATC 2775
Db      2530 TTCACCAACCTCAAGGAACAGGCACGTTCGCTGGGTCTCGAAAGCCGCTGGCACGAGGTG 2589
Qy      2776 GAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCC 2835
Db      2590 GCGCAGGCCTATGCCGATGCCAACAGGATGTTCGGCGATATCGTCAAGGTGACGCCGTCC 2649
Qy      2836 TCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGAC 2895
Db      2650 TCCAAGGTAGTCGGCGACATGGCGCTGATGATGGTGAGCCAGGACCTGACGGTGGCCGAT 2709
Qy      2896 TTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGC 2955
Db      2710 GTCGAAAACCCTGACCGTGAAGTGTCCTTCCCGGACTCGGTGGTATCGATGCTGAAGGGC 2769
Qy      2956 GAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGC 3015
Db      2770 GATCTCGGTCAATCGCCGGGCGGCTGGCCGGAAGCGCTGCAGAAGAAGGCGCTGAAGGGT 2829
Qy      3016 CGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGA----------------AG 3058
Db      2830 GAAAAGCCTTACACGGTTCGTCCGGGTTCGCTTCTGGAAGATGCCGATCTCGATGCCGAA 2889
Qy      3059 AGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGC---- 3114
Db      2890 CGGAAGGTGATCGAGACCAAGCTGGAGCGCAAGGTCGATGACTTCGAGTTTGCCTCCTAT 2949
Qy      3115 CTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACC 3174
Db      2950 CTGATGTATCCGAAGGTGTTCACCGATTTCGCGCTGACCGCCGAGACATACGGCCCGGTC 3009
Qy      3175 TCTGCGCTGGATGATCGTGAATTCTTCTACGGCCTGGTCGAAGGCCGCGAGACTTTGATC 3234
Db      3010 TCCGTGCTGCCCACCCATGCCTATTTCTATGGCATGGAGGATGGCGAAGAG---CTGTTT 3066
Qy      3235 CGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGAC 3294
Db      3067 GCGGATATCGAGCGCGGCAAGACGCTTGTCATCGTCAATCAGGCTTCGTCCGGCATTGAC 3126
Qy      3295 GATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTG 3354
Db      3127 GATAAGGGCATGGTGACGGTGTTCTTCGAGATCAACGGCCAGCCGCGCCGCATCAAGGTG 3186
Qy      3355 CGTGACCGCTCC---GTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAAC 3411
Db      3187 CCGGATCGCGCCCATGGCGCTTCCGGCTCTGCGGTGCGCCGCAAGGCAGAACCCGGCAAT 3246
Qy      3412 AAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCAC---CGTGACTGTTGCTGAAGGT 3468
Db      3247 GCGTCCCATATCGGCGCGCCGATGCCTGGTGTCATCAGCCGCGTCTTCATCAACCAGGGT 3306
Qy      3469 GATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACA 3528
Db      3307 CAGGAGGTCAAGGCCGGCGACGTGCTGCTCTCGATCGAGGCCATGAAGATGGAAACGGCG 3366
Qy      3529 ATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTGGTTCCTGCTGCAACGAAGGTG 3588
Db      3367 CTGCATGCCGAACGCGACGGCAAGATCGCTGAAGTTCTGGTGAAGCCCGGTGACCAGATC 3426
Qy      3589 GAAGGTGGCGACTTGATCGTC 3609
Db      3427 GACGCCAAGGATTTGTTGATC 3447
```

Search completed: March 24, 2004, 06:08:42
Job time : 1227 secs

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 18:35:49 ; Search time 8670 Seconds
(without alignments)
12471.852 Million cell updates/sec

Title: US-10-045-072-1
Perfect score: 3621
Sequence: 1 tggggcggggttagatcctg.....tgatcgtcgtcgtttcctaa 3621

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	609.8	16.8	3537	29	AY408287	AY408287 Mus muscu
2	606.2	16.7	3537	29	AY408285	AY408285 Homo sapi
3	442.4	12.2	3537	29	AY408286	AY408286 Pan trogl
4	432.2	11.9	2877	11	AK077665	AK077665 Mus muscu
c 5	240.6	6.6	888	29	CNS074FW	AL428754 clone BA0
6	236	6.5	903	13	BQ881271	BQ881271 AGENCOURT
c 7	223.4	6.2	859	28	AQ875053	AQ875053 V120F6 mT
8	220.2	6.1	637	10	BF251052	BF251052 EST418309
9	216.6	6.0	862	14	CA793824	CA793824 AGENCOURT
10	216.4	6.0	675	9	AI109467	AI109467 GH08678.5
11	214.8	5.9	744	12	BI103175	BI103175 602889415
12	213.2	5.9	800	14	CK000115	CK000115 AGENCOURT
13	211	5.8	640	12	BI618593	BI618593 RH49459.5
14	208.2	5.7	715	12	BG910468	BG910468 602806142
15	204.4	5.6	641	12	BJ063722	BJ063722 BJ063722
16	202.8	5.6	635	14	CB576991	CB576991 AMGNNUC:C
17	202.8	5.6	696	12	BP177875	BP177875 BP177875
c 18	200.6	5.5	516	28	BZ293357	BZ293357 CG0285.r1
19	200.6	5.5	758	14	CF869970	CF869970 tric021xb
20	200.6	5.5	813	14	CB900127	CB900127 tric021xb
21	199	5.5	887	10	BF781506	BF781506 602104410
22	198.4	5.5	546	14	CB279857	CB279857 ru88d01.y
23	196.6	5.4	515	9	AU202878	AU202878 AU202878
24	195.6	5.4	508	12	BI621845	BI621845 RH53265.5
25	193.6	5.3	666	12	BI605019	BI605019 RH70532.5
26	192.8	5.3	2403	11	AK007782	AK007782 Mus muscu
27	192.8	5.3	2411	11	AK031072	AK031072 Mus muscu
28	191	5.3	625	9	AI293242	AI293242 GH16378.5
29	190.6	5.3	611	10	BE312170	BE312170 601152484
30	189.2	5.2	915	13	BQ730755	BQ730755 AGENCOURT
31	188.4	5.2	1030	29	CNS06M6L	AL405091 T3 end of
32	188.2	5.2	949	13	BQ752438	BQ752438 EST633001
33	188	5.2	921	13	BQ959497	BQ959497 AGENCOURT
34	187.4	5.2	599	10	BF506511	BF506511 AT09003.5
35	185.2	5.1	910	14	CD244370	CD244370 AGENCOURT
36	185	5.1	923	13	BQ734080	BQ734080 AGENCOURT
c 37	184.8	5.1	553	28	AQ501087	AQ501087 V27H5 mTn
38	184.4	5.1	895	10	BF535766	BF535766 602051271
39	184	5.1	616	12	BJ124629	BJ124629 BJ124629
40	183.8	5.1	813	12	BG761035	BG761035 602717545
41	183.6	5.1	556	9	AI134619	AI134619 GH12134.5
42	183.2	5.1	860	12	BI872878	BI872878 603397491
43	182.4	5.0	725	12	BG599744	BG599744 EST504639
44	181.6	5.0	609	9	AI526573	AI526573 uj40b10.y
45	181.2	5.0	610	12	BI082302	BI082302 602877488

ALIGNMENTS

RESULT 1
AY408287
LOCUS AY408287 3537 bp DNA linear GSS 15-DEC-2003
DEFINITION Mus musculus PC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY408287
VERSION AY408287.1 GI:39764258
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3537)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3537)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .3537
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
gene <1. .>3537
/gene="PC"
/locus_tag="HCM3173"
ORIGIN

Query Match 16.8%; Score 609.8; DB 29; Length 3537;
Best Local Similarity 48.5%; Pred. No. 2.2e-131;
Matches 1361; Conservative 0; Mismatches 1420; Indels 24; Gaps 6;

```
Qy        233 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
Db        110 TCAAGAAAGTAATGGTGGCCAACAGAGGTGAGATTGCCATCCGAGTGTTTCGTGCGTGCA 169
Qy        293 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 352
Db        170 CAGAGCTGGGTATCCGCACAGTGGCTGTCTACTCGGAGCAGGACACAGGGCAGATGCACA 229
Qy        353 GCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACC 412
Db        230 GGCAGAAAGCTGATGAAGCCTACCTTATTGGCCGTGGCCTGGCACCTGTGCAGGCCTACC 289
Qy        413 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGG 472
Db        290 TGCACATTCCAGACATCATCAAGGTGGCCAAGGAAAATGGTGTAGATGCGGTGCATCCTG 349
Qy        473 GATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTA 532
Db        350 GCTATGGGTTCCTCTCAGAGCGAGCAGACTTTGCCCAGGCCTGCCAAGATGCTGGAGTCC 409
Qy        533 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAA 592
Db        410 GGTTCATTGGTCCAAGCCCAGAGGTGGTCCGCAAGATGGGAGACAAGGTGGAAGCCCGGG 469
Qy        593 CCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATC---CACCCCGAGCAAAAACA 649
Db        470 CCATTGCCATCGCTGCAGGCGTTCCAGTGGTCCCCGGCACGGACTCCCCCATCAGCTCCC 529
Qy        650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
Db        530 TGCACGAGGCGCATGAGTTCTCCAACACCTACGGCTTCCCTATTATCTTCAAGGCCGCCT 589
Qy        710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
Db        590 ACGGAGGTGGGGGCCGCGGCATGCGGGTCGTGCATAGCTATGAGGAGTTGGAAGAGAATT 649
Qy        770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
Db        650 ACACCCGGGCCTACTCCGAGGCCTTGGCAGCCTTTGGGAATGGAGCGTTGTTTGTGGAGA 709
Qy        830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
Db        710 AGTTCATTGAGAAGCCAAGGCACATTGAGGTGCAGATCCTAGGGGACCAGTATGGGAACA 769
Qy        890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
Db        770 TCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGTCGGCACCAGAAGGTGGTAGAGA 829
Qy        950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
Db        830 TCGCCCCTGCTACCCACCTGGATCCCCAACTTCGCTCACGTCTCACCAGTGACTCTGTCA 889
Qy       1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
Db        890 AACTTGCCAAGCAGGTAGGCTATGAGAACGCCGGCACTGTGGAGTTCCTGGTGGACAAGC 949
```

```
Qy       1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
Db        950 ACGGCAAGCACTACTTCATCGAGGTCAATTCCCGCCTGCAGGTGGAGCACACGGTCACCG 1009
Qy       1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
Db       1010 AGGAGATCACAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1069
Qy       1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
Db       1070 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1129
Qy       1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
Db       1130 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTTT 1189
Qy       1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAA 1369
Db       1190 TCCGGAGTGGTGAGGGCATGGGCATCCGCCTGGACAACGCCTCTGCATTCCAGGGCGCTG 1249
Qy       1370 TCA---CCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTG 1426
Db       1250 TCATATCGCCCCACTATGACTCTCTGCTCGTCAAGGTCATTGCACACGGCAAAGACCACC 1309
Qy       1427 AAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAA 1486
Db       1310 CCACAGCTGCCACCAAGATGAGCAGAGCCCTGGCCGAGTTCCGTGTCCGAGGTGTAAAGN 1369
Qy       1487 CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCG 1546
Db       1370 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1429
Qy       1547 CCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGC 1606
Db       1430 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1489
Qy       1607 AGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTC 1666
Db       1490 NNNNNNNNNNNNNNNNNNNNNNNNGACATGTCATGGTGAATGGCCCTACCACTCCAATCC 1549
Qy       1667 CAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATC------TGCCAC 1720
Db       1550 CTGTCAATGTGAGCCCCAGTCCTGTGGATCCTGCTGTTCCTGTGGTGCCCATAGGCCCAC 1609
Qy       1721 TGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC 1780
Db       1610 CTCCAGCTGGTTTCAGGGACATCCTTCTGCGAGAAGGGCCAGAGGGCTTTGCCCGAGCTG 1669
Qy       1781 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTT 1840
Db       1670 TGCGGAATCACCAGGGGCTGCTGTTGATGGACACAACCTTCCGGGATGCCCACCAGTCAC 1729
Qy       1841 TGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGC 1900
Db       1730 TACTGGCCACTAGAGTGCGCACACATGATCTCAAAAAGATTGCGCCCTATGTTGCCCACA 1789
Qy       1901 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGC 1960
Db       1790 ACTTCAACAAGCTCTTCAGCATGGAGAACTGGGGAGGCGCCACGTTCGACGTTGCCATGC 1849
Qy       1961 GTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 2020
Db       1850 GCTTCCTGTACGAGTGCCCCTGGCGGCGGCTCCAGGAGCTCCGGGAGCTTATCCCGAACA 1909
Qy       2021 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080
Db       1910 TCCCGTTCCAGATGCTACTGAGGGGGGCCAATGCTGTGGGCTACACCAACTACCCTGACA 1969
Qy       2081 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 2140
Db       1970 ACGTGGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGAATGGTATGGACGTCTTCCGAGTCT 2029
```

```
Qy      2141 TCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCA 2200
Db      2030 TTGACTCCCTCAACTACTTGCCAAACATGCTGCTGGGCATGGAAGCAG---CAGGCAGTG 2086

Qy      2201 ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAA 2260
Db      2087 CTGGGGGTGTGGTGGAGGCTGCCATCTCATACACGGGGGACGTGGCTGACCCTAGTCGCA 2146

Qy      2261 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2320
Db      2147 CTAAATACTCACTGGAGTACTACATGGGCTTAGCTGAAGAACTGGTGCGAGCTGGCACTC 2206

Qy      2321 ACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2380
Db      2207 ACATCCTGTGCATTAAGGACATGGCGGGCCTGCTGAAGCCTGCCGCCTGCACCATGCTGG 2266

Qy      2381 TCACCGCACTGCGCCGTGAATT---CGATCTGCCAGTGCACGTGCACACCCACGACACTG 2437
Db      2267 TCAGCTCCCTCCGGGACCGATTCCCCGACCTCCCACTGCACATCCATACCCATGATACAT 2326

Qy      2438 CGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497
Db      2327 CAGGGGCAGGTGTGGCAGCCATGCTGGCCTGTGCACAAGCAGGGGCTGATGTTGTGGACG 2386

Qy      2498 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTG 2557
Db      2387 TGGCAGTAGACTCCATGTCTGGGATGACCTCACAGCCAAGCATGGGGGCCCTGGTGGCCT 2446

Qy      2558 CATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGC 2617
Db      2447 GTACCAAAGGGACTCCTTTGGACACAGAGGTACCCCTGGAGCGTGTGTTTGACTACAGTG 2506

Qy      2618 CGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGA------GTCTGGAACCCCAG 2671
Db      2507 AGTACTGGGAAGGGGCTCGGGGACTGTACGCAGCCTTCGATTGCACGGCTACCATGAAGT 2566

Qy      2672 GCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTG 2731
Db      2567 CTGGCAACTCCGACGTGTATGAGAATGAGATTCCAGGGGGCCAGTACACCAACCTGCACT 2626

Qy      2732 CACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAG 2791
Db      2627 TCCAGGCCCATAGCATGGGGCTTGGCTCCAAGTTCAAGGAGGTCAAGAAGGCCTATGTGG 2686

Qy      2792 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCG 2851
Db      2687 AGGCTAACCAGATGCTGGGGGACCTCATCAAGGTGACACCATCCTCCAAGATTGTGGGGG 2746

Qy      2852 ACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCAC 2911
Db      2747 ACCTGGCCCAGTTCATGGTGCAGAATGGGTTGAGCCGGGCAGAGGCAGAAGCTCAGGCAG 2806

Qy      2912 AAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTC 2971
Db      2807 AAGAGCTGTCCTTCCCCCGCTCTGTGGTGGAGTTCCTGCAGGGCTACATTGGCATTCCCC 2866

Qy      2972 CAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCC 3016
Db      2867 ATGGGGGTTTCCCTGAGCCCTTTCGCTCTAAGGTGCTAAAGGACC 2911
```

RESULT 2
AY408285

```
LOCUS       AY408285               3537 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens PC gene, VIRTUAL TRANSCRIPT, partial sequence, genomic
            survey sequence.
ACCESSION   AY408285
VERSION     AY408285.1  GI:39764256
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3537)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3537)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3537
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>3537
                     /gene="PC"
                     /locus_tag="HCM3173"
ORIGIN
```

Query Match 16.7%; Score 606.2; DB 29; Length 3537;
Best Local Similarity 48.7%; Pred. No. 1.5e-130;
Matches 1365; Conservative 0; Mismatches 1416; Indels 24; Gaps 7;

```
Qy       233 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
Db       110 TCAAGAAAGTCATGGTGGCCAACAGAGGTGAGATTGCCATCCGTGTGTTCCGGGCCTGCA 169

Qy       293 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 352
Db       170 CGGAGCTGGGCATCCGCACCGTAGCCATCTACTCTGAGCAGGACACGGGCCAGATGCACC 229

Qy       353 GCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACC 412
Db       230 GGCAGAAAGCAGATGAAGCCTATCTCATCGGCCGCGGCCTGGCCCCCGTGCAGGCCTACC 289

Qy       413 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGG 472
Db       290 TGCACATCCCAGACATCATCAAGGTGGCCAAGGAGAACAACGTAGATGCAGTGCACCCTG 349

Qy       473 GATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTA 532
Db       350 GCTACGGGTTCCTCTCTGAGCGAGCGGACTTCGCCCAGGCCTGCCAGGATGCAGGGGTCC 409

Qy       533 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAA 592
Db       410 GGTTTATTGGGCCAAGCCCAGAAGTGGTCCGCAAGATGGGAGACAAGGTGGAGGCCCGGG 469

Qy       593 CCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGG---AATCCACCCCGAGCAAAAACA 649
Db       470 CCATCGCCATTGCTGCGGGTGTTCCCGTTGTCCCTGGCACAGATGCCCCCATCACGTCCC 529

Qy       650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
Db       530 TGCATGAGGCCCACGAGTTCTCCAACACCTACGGCTTCCCCATCATCTTCAAGGCGGCCT 589

Qy       710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
Db       590 ATGGGGGTGGAGGGCGTGGCATGAGGGTGGTGCACAGCTACGAGGAGCTGGAGGAGAATT 649

Qy       770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
Db       650 ACACCCGGGCCTACTCAGAGGCTCTGGCCGCCTTTGGGAATGGGGCGCTGTTTGTGGAGA 709
```

```
Qy       830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
Db       710 AGTTCATCGAGAAGCCACGGCACATCGAGGTGCAGATCTTGGGGGACCAGTATGGGAACA 769

Qy       890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
Db       770 TCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGGCGGCACCAGAAGGTGGTCGAGA 829

Qy       950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
Db       830 TTGCCCCCGCCGCCCACCTGGACCCGCAGCTTCGGACTCGGCTCACCAGCGACTCTGTGA 889

Qy      1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
Db       890 AACTCGCTAAACAGGTGGGCTACGAGAACGCAGGCACCGTGGAGTTCCTGGTGGACAGGC 949

Qy      1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
Db       950 ACGGCAAGCACTACTTCATCGAGGTCAACTCCCGCCTGCAGGTGGAGCACACGGTCACAG 1009

Qy      1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
Db      1010 AGGAGATCACCGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1069

Qy      1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
Db      1070 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1129

Qy      1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
Db      1130 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGT 1189

Qy      1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1366
Db      1190 TCCGGAGCGGAGAGGGCATGGGCATCCGCCTGGATAATGCTTCCGCCTTCCAAGGAGCCG 1249

Qy      1367 AAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTG 1426
Db      1250 TCATCTCGCCCCACTACGACTCCCTGCTGGTCAAAGTCATTGCCCACGGCAAAGACCACC 1309

Qy      1427 AAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAA 1486
Db      1310 CCACGGCCGCCACCAAGATGAGCAGGGCCCTTGCGGAGTTCCGCGTCCGAGGTGTGAAGN 1369

Qy      1487 CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCG 1546
Db      1370 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1429

Qy      1547 CCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGATGAGC 1606
Db      1430 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1489

Qy      1607 AGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAA----CAAGCCTCATGGTGTG 1662
Db      1490 NNNNNNNNNNNNNNNNNNNNNNNGCCATGTCATGGTAAACGGTCCAACCACCCCGATTC 1549

Qy      1663 CGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG--CCTAACATCAAGGATCTGCCAC 1720
Db      1550 CCGTCAAGGCCAGCCCCAGCCCCACGGACCCCGTTGTCCCTGCAGTGCCCATAGGCCCGC 1609

Qy      1721 TGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC 1780
Db      1610 CCCCGGCTGGTTTCAGAGACATCCTGCTGCGAGAGGGGCCTGAGGGCTTTGCTCGAGCTG 1669

Qy      1781 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTT 1840
Db      1670 TGCGGAACCACCCGGGGCTGCTGCTGATGGACACGACCTTCAGGGACGCCCACCAGTCAC 1729

Qy      1841 TGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGC 1900
Db      1730 TGCTGGCCACTCGTGTGCGCACCCACGATCTCAAAAAGATCGCCCCCTATGTTGCCCACA 1789

Qy      1901 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGC 1960
```

```
Db      1790 ACTTCAGCAAGCTCTTCAGCATGGAGAACTGGGGAGGAGCCACGTTTGACGTCGCCATGC 1849

Qy      1961 GTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 2020
Db      1850 GCTTCCTGTATGAGTGCCCCTGGCGGCGGCTGCAGGAGCTCCGGGAGCTCATCCCCAACA 1909

Qy      2021 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080
Db      1910 TCCCTTTCCAGATGCTGCTGCGGGGGGCCAATGCTGTGGGCTACACCAACTACCCAGACA 1969

Qy      2081 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 2140
Db      1970 ACGTGGTCTTCAAGTTCTGTGAAGTGGCCAAAGAGAATGGCATGGATGTCTTCCGTGTGT 2029

Qy      2141 TCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCA 2200
Db      2030 TTGACTCCCTCAACTACTTGCCCAACATGCTGCTGGGCATGGAGGCGG---CAGGAAGTG 2086

Qy      2201 ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAA 2260
Db      2087 CCGGAGGCGTGGTGGAGGCTGCCATCTCATACACGGGCGACGTGGCCGACCCCAGCCGCA 2146

Qy      2261 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2320
Db      2147 CCAAGTACTCACTGCAGTACTACATGGGCTTGGCCGAAGAGCTGGTGCGAGCTGGCACCC 2206

Qy      2321 ACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2380
Db      2207 ACATCCTGTGCATCAAGGACATGGCCGGGCTGCTGAAGCCCACGGCCTGCACCATGCTGG 2266

Qy      2381 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTG 2437
Db      2267 TCAGCTCCCTCCGGGACCGCTTCCCCGACCTCCCACTGCACATCCACACCCACGACACGT 2326

Qy      2438 CGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497
Db      2327 CAGGGGCAGGCGTGGCAGCCATGCTGGCCTGTGCCCAGGCTGGAGCTGATGTGGTGGATG 2386

Qy      2498 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTG 2557
Db      2387 TGGCAGCTGATTCCATGTCTGGGATGACTTCACAGCCCAGCATGGGGGCCCTGGTGGCCT 2446

Qy      2558 CATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGC 2617
Db      2447 GTACCAGAGGGACTCCCCTGGACACAGAGGTGCCCATGGAGCGCGTGTTTGACTACAGTG 2506

Qy      2618 CGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAA 2677
Db      2507 AGTACTGGGAGGGGGCTCGGGGACTGTACGCGGCCTTCGACTGCACGGCCACCATGAAGT 2566

Qy      2678 CCGG------TCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTG 2731
Db      2567 CTGGCAACTCGGACGTGTATGAAAATGAGATCCCAGGGGGCCAGTACACCAACCTGCACT 2626

Qy      2732 CACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACTCATCGAAGACAACTACGCAG 2791
Db      2627 TCCAGGCCCACAGCATGGGGCTTGGCTCCAAGTTCAAGGAGGTCAAGAAGGCCTATGTGG 2686

Qy      2792 CCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCG 2851
Db      2687 AGGCCAACCAGATGCTGGGCGATCTCATCAAGGTGACGCCCTCCTCCAAGATCGTGGGGG 2746

Qy      2852 ACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCAC 2911
Db      2747 ACCTGGCCCAGTTTATGGTGCAGAATGGATTGAGCCGGGCAGAGGCCGAAGCTCAGGCGG 2806

Qy      2912 AAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTC 2971
Db      2807 AAGAGCTGTCCTTTCCCCGCTCCGTGGTGGAGTTCCTGCAGGGCTACATCGGTGTCCCCC 2866

Qy      2972 CAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCC 3016
```

```
Db       2867 ATGGGGGGTTCCCCGAACCCTTTCGCTCTAAGGTACTGAAGGACC 2911


RESULT 3
AY408286
LOCUS       AY408286               3537 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Pan troglodytes PC gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY408286
VERSION     AY408286.1  GI:39764257
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 3537)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3537)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3537
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>3537
                     /gene="PC"
                     /locus_tag="HCM3173"
ORIGIN

  Query Match             12.2%;  Score 442.4;  DB 29;  Length 3537;
  Best Local Similarity   41.5%;  Pred. No. 4e-92;
  Matches 1004;  Conservative    0;  Mismatches 1400;  Indels   18;  Gaps    6;

Qy        233 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCAC 292
Db        110 TCAAGANNGTCATGGTGGCCAACAGAGGTGAGATTGCCATCCGTGTGTTCCGGGCCTGCA 169

Qy        293 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 352
Db        170 CGGAGCTGGGCATCCGCACCGTAGCCATCTACTCTGAGCAGGACACGGGCCAGATGCACC 229

Qy        353 GCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACC 412
Db        230 GGCAGAAAGCAGATGAAGCCTATCTCATCGGCCGCGGCCTGGCCCCCGTGCAGGCCTACC 289

Qy        413 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCCATTTACCCGG 472
Db        290 TGCACATCCCAGACATCATCAAGGTGGCCAAGGAGAACAACGTAGATGCAGTACACCCTG 349

Qy        473 GATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTA 532
Db        350 GCTACGGGTTCCTCTCTGAGCGAGCGGACTTCGCCCAGGCCTGCCAGNATGCAGGGGTCC 409

Qy        533 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAA 592
Db        410 GGTTTATTGGGCCAAGCCCAGAAGTGGTCCGCAAGATGGGAGACAAGGTGGAGGCCCGGG 469

Qy        593 CCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGG---AATCCACCCCGAGCAAAAACA 649

Qy        650 TCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTG 709
Db        530 TGCATGAGGCCCACGAGTTCTCCAACACCTACGGCTTCCCCATCATCTTCAAGGCGGCCT 589

Qy        710 CCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAG 769
Db        590 ATGGGGGTGGAGGGCGTGGCATGAGGGTGGTGCACAGCTACGAGNNNNNNNNNNNNNNNN 649

Qy        770 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAAC 829
Db        650 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 709

Qy        830 GTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAG 889
Db        710 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 769

Qy        890 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAA 949
Db        770 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATCCAGCGGCGGCACCAGAAGNNNNTCGAGA 829

Qy        950 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAA 1009
Db        830 TTGCCCCCGCCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 889

Qy       1010 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 1069
Db        890 NNNNNNNNNNNNNNGTGGGCTACGAGAACGCAGGCACCGTGGAGTTCCTGGTGGACAGGC 949

Qy       1070 AGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 1129
Db        950 ACGGCAAGCACTACTTCATCGAGGTCAACTCCCGCCTGCAGGTGGAGCACACGGTCACAG 1009

Qy       1130 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1189
Db       1010 AGGAGATCACCGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1069

Qy       1190 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1249
Db       1070 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1129

Qy       1250 GCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1309
Db       1130 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1189

Qy       1310 ACCGCTCACCAGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCG 1366
Db       1190 NNNGGAGTGGAGAGGGCATGGGCATNNGCCTGGATAATNCTTCCGCCTTCCAAGGAGCCG 1249

Qy       1367 AAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTG 1426
Db       1250 TCANCTCGCCCCACTACGACTCCCTGCTGGTCAAAGTCATTGCCCACGGCAAAGACCACC 1309

Qy       1427 AAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGT-----GTCTGGTGT 1481
Db       1310 CCACGGCCGCCACCAAGATGAGCAGGGCCCTTGCGGAGTTCCGCGTCCGAGGTGTGAAGN 1369

Qy       1482 TGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCG 1541
Db       1370 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1429

Qy       1542 CATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGA 1601
Db       1430 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1489

Qy       1602 TGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGT 1661
Db       1490 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACCACCCCGATTC 1549

Qy       1662 GCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACT 1721
```

```
Db       1550 CCGTCAAGGCCNGNNNNNNNCNCACGGACCCCATTGTCCCTGCAGTGCCTATAGGCCCGC 1609

Qy       1722 GCCACGC-GGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC 1780

Db       1610 CCCCGGCTGGTTTCAGAGACATCCTGCTGCGAGAGGGGCCTGAGGGCTTTGCTCGAGCTG 1669

Qy       1781 TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTT 1840

Db       1670 TGCGGAACCACCCGGGGCTGCTGCTGATGGACACGACCTTCAGGGACGCCCACCAGTCAC 1729

Qy       1841 TGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGC 1900

Db       1730 TGCTGGCCACTCGTGTGCGCACCCACGATCTCAAAAAGATCGCCCCCTATGTTGCCCACA 1789

Qy       1901 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGC 1960

Db       1790 ACTTCAGCAAGCTCTTCAGCATGGAGAANNNNNNNNGAGCCACGTTTGACGTCGCCATGC 1849

Qy       1961 GTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 2020

Db       1850 GCTTCCTGTATGAGTGCCCCTGGCGGCGGCTGCAGGAGCTCCGGGAGCTCATCCCCAACA 1909

Qy       2021 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGACT 2080

Db       1910 TCCCTTTCCAGATGCTGCTGCGGGGGGCCAATGCTGTGGGCTACACCAACTACCCAGACA 1969

Qy       2081 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 2140

Db       1970 ACGTGGTCTTCAAGTTCTGTGAAGTGGCCAAAGAGAATGGCATGGATGTCTTCCGTGTGT 2029

Qy       2141 TCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCA 2200

Db       2030 TTGACTCCCTCAACTACTTGCCCAACATGCTGCTGGGCATGGAGGCGG---CAGGAAGTG 2086

Qy       2201 ACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAA 2260

Db       2087 CCGGAGGCGTGGTGGAGGCTGCCATCTCATACACGGGCGACGTGGCCGACCCCAGCCGCA 2146

Qy       2261 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTC 2320

Db       2147 CCAAGTACTNACTGCAGTACTACATGGGCTTGGCCGAANNGCTGGTGCGAGCTGGCACCC 2206

Qy       2321 ACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2380

Db       2207 ACATCCTGTGCATCAAGGACATGGCCGGGCTGCTGAAGCCCACGGCCTGCACCATGCTGG 2266

Qy       2381 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTG 2437

Db       2267 TCAGCTCCCTCCGGGACCGCTTCCCCGACCTCCCACTGCACATCCACACCCACGACACGT 2326

Qy       2438 CGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2497

Db       2327 CAGGGGCAGGCGTGGCAGCCATGCTGGCCTGTGCCCAGGCTGGAGCTGATGTGGTGGATG 2386

Qy       2498 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTG 2557

Db       2387 TGGCAGCTGATTCCATGTCTGGGATGACTTCACAGCCCAGCATGGGGGCCCTGGTGGCCT 2446

Qy       2558 CATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGC 2617

Db       2447 GTACCAGAGGGACTCCCCTGGACACAGAGGTGCCCATGGAGCGCGTGTTTGACTACAGCG 2506

Qy       2618 CGTACTGGGAAGCAGTGCGCGG 2639

Db       2507 AGTACTGGGAGGGGGCTCGGGG 2528
```

RESULT 4
AK077665
LOCUS AK077665 2877 bp mRNA linear HTC 18-SEP-2003
DEFINITION Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730513C12 product:pyruvate decarboxylase, full insert sequence.
ACCESSION AK077665
VERSION AK077665.1 GI:26346580
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2877)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

```
          URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .2877
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:5730513C12"
                     /db_xref="MGI:2393843"
                     /db_xref="taxon:10090"
                     /clone="5730513C12"
                     /tissue_type="whole body"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="8 days embryo"
     CDS             <1. .2489
                     /note="unnamed protein product; putative
                     pyruvate decarboxylase (MGD|MGI:97520, GB|NM_008797,
                     evidence: BLASTN, 99%, match=2877)"
                     /codon_start=3
                     /protein_id="BAC36939.1"
                     /db_xref="GI:26346581"
                     /translation="VSEGQSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRI
                     EVFRSGEGMGIRLDNASAFQGAVISPHYDSLLVKVIAHGKDHPTAATKMSRALAEFRV
                     RGVKTNIPFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQKLLHYLGHVMVN
                     GPTTPIPVNVSPSPVDPAVPVVPIGPPPAGFRDILLREGPEGFARAVRNHQGLLLMDT
                     TFRDAHQSLLATRVRTHDLKKIAPYVAHNFNKLFSMENWGGATFDVAMRFLYECPWRR
                     LQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLP
                     NMLQGMEAAGSAGGVVEAAISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKD
                     MAGLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVAVDS
                     MSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEGARGLYAAFDCTATMKSGN
                     SDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGD
                     LAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLPRI
                     EGRPGASLPPLNLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLDS
                     LNTRLFLQGPKIAEEFEVELERGKTLHIKALAVSDLNRAGQRQVFFELNGQLRSILVK
                     DTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDIKVAAGDKVAKGQPLCVLSAMKMETV
                     VTSPMEGTIRKVHVTKDMTLEGDDLILEIE"
     polyA_signal    2851. .2856
                     /note="putative"
     polyA_site      2877
                     /note="putative"
ORIGIN

  Query Match             11.9%;  Score 432.2;  DB 11;  Length 2877;
  Best Local Similarity   53.9%;  Pred. No. 8.9e-90;
  Matches 1006;  Conservative    0;  Mismatches  838;  Indels   21;  Gaps    5;

Qy       1170 CTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCA 1229
Db          2 CGTGTCCGAAGGCCAGAGCCTGCCTGACCTGGGCCTGCGGCAGGAGAACATCCGCATCAA 61

Qy       1230 CGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGA 1289
Db          62 TGGCTGTGCCATTCAGTGTCGGGTCACCACCGAGGACCCTGCACGCAGCTTCCAGCCAGA 121

Qy       1290 TACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGC 1349
Db         122 CACCGGCCGCATTGAGGTTTTCCGGAGTGGTGAGGGCATGGGCATCCGCCTGGACAACGC 181

Qy       1350 AGCTCAGCTCGGTGGCGAAATCA---CCGCACACTTTGACTCCATGCTGGTGAAAATGAC 1406
Db         182 CTCTGCATTCCAGGGCGCTGTCATATCGCCCCACTATGACTCTCTGCTCGTCAAGGTCAT 241

Qy       1407 CTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTT 1466
Db         242 TGCACACGGCAAAGACCACCCCACAGCTGCCACCAAGATGAGCAGAGCCCTGGCCGAGTT 301

Qy       1467 CACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGA 1526
Db         302 CCGTGTCCGAGGTGTAAAGACCAACATCCCCTTCCTGCAGAATGTTCTCAACAACCAGCA 361

Qy       1527 CTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGC 1586
Db         362 GTTCCTGGCAGGCACAGTGGACACCCAGTTCATCGATGAGAACCCTGAGCTGTTCCAGCT 421

Qy       1587 TCCACCTGCTGATGATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAA 1646
Db         422 TCGGCCTGCACAGAACCGGGCCCAGAAGTTGCTACATTACCTCGGACATGTCATGGTGAA 481

Qy       1647 CAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACAT 1706
Db         482 TGGCCCTACCACTCCAATCCCTGTCAATGTGAGCCCCAGTCCTGTGGATCCTGCTGTTCC 541

Qy       1707 CAAGGATC------TGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCC 1760
Db         542 TGTGGTGCCCATAGGCCCACCTCCAGCTGGTTTCAGGGACATCCTTCTGCGAGAAGGGCC 601

Qy       1761 AGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTT 1820
Db         602 AGAGGGCTTTGCCCGAGCTGTGCGGAATCACCAGGGGCTGCTGTTGATGGACACAACCTT 661

Qy       1821 CCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGC 1880
Db         662 CCGGGATGCCCACCAGTCACTACTGGCCACTAGAGTGCGCACACATGATCTCAAAAAGAT 721

Qy       1881 GGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGC 1940
Db         722 TGCGCCCTATGTTGCCCACAACTTCAACAAGCTCTTCAGCATGGAGAACTGGGGAGGCGC 781

Qy       1941 GACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCT 2000
Db         782 CACGTTCGACGTTGCCATGCGCTTCCTGTACGAGTGCCCCTGGCGGCGGCTCCAGGAGCT 841

Qy       2001 GCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGG 2060
Db         842 CCGGGAGCTTATCCCGAACATCCCGTTCCAGATGCTACTGAGGGGGGCCAATGCTGTGGG 901

Qy       2061 ATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGG 2120
Db         902 CTACACCAACTACCCTGACAACGTGGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGAATGG 961

Qy       2121 CGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAAT 2180
Db         962 TATGGACGTCTTCCGAGTCTTTGACTCCCTCAACTACTTGCCAAACATGCTGCAGGGCAT 1021

Qy       2181 CGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGA 2240
Db         1022 GGAAGCAG---CAGGCAGTGCTGGGGGTGTGGTGGAGGCTGCCATCTCATACACGGGGGA 1078

Qy       2241 TCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGA 2300
Db         1079 CGTGGCTGACCCTAGTCGCACTAAATACTCACTGGAGTACTACATGGGCTTAGCTGAAGA 1138

Qy       2301 GATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCC 2360
Db         1139 ACTGGTGCGAGCTGGCACTCACATCCTGTGCATTAAGGACATGGCGGGCCTGCTGAAGCC 1198

Qy       2361 AGCTGCGGTAACCAAGCTGGTCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCA 2417
Db         1199 TGCCGCCTGCACCATGCTGGTCAGCTCCCTCCGGGACCGATTCCCCGACCTCCCACTGCA 1258

Qy       2418 CGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGC 2477
Db         1259 CATCCATACCCATGATACATCAGGGGCAGGTGTGGCAGCCATGCTGGCCTGTGCACAAGC 1318

Qy       2478 TGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATC 2537
Db         1319 AGGGGCTGATGTTGTGGACGTGGCAGTAGACTCCATGTCTGGGATGACCTCACAGCCAAG 1378

Qy       2538 CCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATACCGGTTTGAGCCTCGA 2597
Db         1379 CATGGGGGCCCTGGTGGCCTGTACCAAAGGGACTCCTTTGGACACAGAGGTACCCCTGGA 1438

Qy       2598 GGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGA 2657
Db         1439 GCGTGTGTTTGACTACAGTGAGTACTGGGAAGGGGCTCGGGGACTGTACGCAGCCTTCGA 1498

Qy       2658 ------GTCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGG 2711
```

```
Db       1499 TTGCACGGCTACCATGAAGTCTGGCAACTCCGACGTGTATGAGAATGAGATTCCAGGGGG 1558

Qy       2712 ACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGTTTCGAACT 2771

Db       1559 CCAGTACACCAACCTGCACTTCCAGGCCCATAGCATGGGGCTTGGCTCCAAGTTCAAGGA 1618

Qy       2772 CATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCC 2831

Db       1619 GGTCAAGAAGGCCTATGTGGAGGCTAACCAGATGCTGGGGGACCTCATCAAGGTGACACC 1678

Qy       2832 ATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGC 2891

Db       1679 ATCCTCCAAGATTGTGGGGGACCTGGCCCAGTTCATGGTGCAGAATGGGTTGAGCCGGGC 1738

Qy       2892 AGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCG 2951

Db       1739 AGAGGCAGAAGCTCAGGCAGAAGAGCTGTCCTTCCCCCGCTCTGTGGTGGAGTTCCTGCA 1798

Qy       2952 CGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGA 3011

Db       1799 GGGCTACATTGGCATTCCCCATGGGGGTTTCCCTGAGCCCTTTCGCTCTAAGGTGCTAAA 1858

Qy       3012 AGGCC 3016

Db       1859 GGACC 1863
```

```
RESULT 5
CNS074FW/c
LOCUS       CNS074FW                 888 bp    DNA     linear   GSS 07-JUL-2001
DEFINITION  clone BA0AB027F04 of library BA0AB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL428754
VERSION     AL428754.1  GI:12211948
KEYWORDS    GSS.
SOURCE      Kluyveromyces lactis
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 888)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
  MEDLINE   20584721
   PUBMED   11152886
REFERENCE   3  (bases 1 to 888)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Kluyveromyces lactis"
                     /mol_type="genomic DNA"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB027F04"
                     /clone_lib="BA0AB"
     misc_feature    complement(<2. .>886)
                     /note="similar to Saccharomyces cerevisiae ORF YBR218c [
                     PYC2 ; pyruvate carboxylase 2 ]
                     similar to Saccharomyces cerevisiae ORF YGL062w [ PYC1 ;
                     pyruvate carboxylase 1 ]"
                     /evidence=not_experimental
ORIGIN
```

Query Match 6.6%; Score 240.6; DB 29; Length 888;
Best Local Similarity 55.9%; Pred. No. 4.3e-45;
Matches 478; Conservative 0; Mismatches 374; Indels 3; Gaps 1;

```
Qy        395 CACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAG 454

Db        855 CTCCAGTTGGTGCTTACTTGGCGATTGATGAAATTATCAACATTGCCAAGTCCCATGGTG 796

Qy        455 CAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGT 514

Db        795 TGAACTTTATCCATCCTGGTTACGGGTTCCTATCAGAAAACTCTGAATTCGCTGAAAAAG 736

Qy        515 GTGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTG 574

Db        735 TTGCAGCTTCTGGTATCACTTGGGTTGGTCCTCCAGCTGCTGTTATCGATTCCGTTGGTG 676

Qy        575 ATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCA 634

Db        675 ACAAAGTTTCAGCCAGAAACTTGGCTGAAAAGGCTAACGTTCCAGTGGTTCCAGGTACTC 616

Qy        635 C---CCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCA 691

Db        615 CAGGTCCAATTGAGACTGTCGAAGAAGCACAGGCCTTCGTTGATAAATACGGATTCCCAG 556

Qy        692 TCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTG 751

Db        555 TCATCATTAAGGCCGCCTTTGGTGGTGGTGGTCGTGGTATGAGAGTTGTCCGTGAAGGTG 496

Qy        752 ATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGCGATG 811

Db        495 ATGATATCAGTGACGCTTTCCAACGTGCCAAATCTGAAGCTATCACTGCTTTTGGTAACG 436

Qy        812 GCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTG 871

Db        435 GTACTTGTTTCATTGAAAGATTCTTGGACAAGCCAAAGCATATCGAAGTTCAATTGTTGG 376

Qy        872 GCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTC 931

Db        375 CTGATAACTACGGTAACGTTGTTCATCTGTTCGAAAGAGATTGTTCCGTTCAAAGAAGAC 316

Qy        932 ACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCA 991

Db        315 ATCAAAAGGTTGTCGAAGTTGCCCCAGCCAAAACTCTACCTGAAGAAGTGCGTGACGCTA 256

Qy        992 TTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGG 1051

Db        255 TCTTAACTGATGCGGTCAAATTAGCTAAGGAAGCTGGTTACAGAAACGCTGGTACCGCTG 196

Qy       1052 AATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGG 1111

Db        195 AATTCTTAGTCGATAACCAAAACAGACATTACTTTATCGAAATTAACCCAAGAATTCAAG 136
```

```
Qy       1112 TTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCT 1171
Db        135 TCGAACATACAATTACCGAAGAAATCACCGGTATCGATATCGTCGCTGCTCAAATCCAAA 76

Qy       1172 TGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACG 1231
Db          75 TCGCTGCAGGTGCCACTTTGGAACAATTGGGTTTGATGCAAGATAAGATTACAACTCGTG 16

Qy       1232 GTGCAGCACTGCAGT 1246
Db          15 GTTTCGCCATCCAGT 1
```

RESULT 6
BQ881271
LOCUS BQ881271 903 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_7978339 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6185002 5', mRNA sequence.
ACCESSION BQ881271
VERSION BQ881271.1 GI:22273279
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13575 row: b column: 11
High quality sequence stop: 692.
FEATURES Location/Qualifiers
source 1. .903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185002"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
ORIGIN

Query Match 6.5%; Score 236; DB 13; Length 903;
Best Local Similarity 57.4%; Pred. No. 5.2e-44;
Matches 466; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

```
Qy       1743 CCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGC 1802
Db          18 CCTGCTGCGAGAGGGGCCTGAGGGCTTTGCTCGAGCTGTGCGGAACCACCCGGGGCTGCT 77

Qy       1803 AGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTC 1862
```

```
Db          78 GCTGATGGACACGACCTTCAGGGACGCCCACCAGTCACTGCTGGCCACTCGTGTGCGCAC 137

Qy       1863 ATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGT 1922
Db         138 CCACGATCTCAAAAAGATCGCCCCCTATGTTGCCCACAACTTCAGCAAGCTCTTCAGCAT 197

Qy       1923 GGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTG 1982
Db         198 GGAGAACTGGGGAGGAGCCACGTTTGACGTCGCCATGCGCTTCCTGTATGAGTGCCCCTG 257

Qy       1983 GGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCG 2042
Db         258 GCGGCGGCTGCAGGAGCTCCGGGAGCTCATCCCCAACATCCCTTTCCAGATGCTGCTGCG 317

Qy       2043 CGGCCGCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAA 2102
Db         318 GGGGGCCAATGCTGTGGGCTACACCAACTACCCAGACAACGTGGTCTTCAAGTTCTGTGA 377

Qy       2103 GGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTC 2162
Db         378 AGTGGCCAAAGAGAATGGCATGGATGTCTTCCGTGTGTTTGACTCCCTCAACTACTTGCC 437

Qy       2163 CCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGC 2222
Db         438 CAACATGCTGCTGGGCATGGAGGCGG---CAGGAAGTGCCGGAGGCGTGGTGGAGGCTGC 494

Qy       2223 TATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTA 2282
Db         495 CATCTCATACACGGGCGACGTGGCCGACCCCAGCCGCACCAAGTACTCACTGCAGTACTA 554

Qy       2283 CCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATAT 2342
Db         555 CATGGGCTTGGCCGAAGAGCTGGTGCGAGCTGGCACCCACATCCTGTGCATCAAGGACAT 614

Qy       2343 GGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCG---CCGTGA 2399
Db         615 GGCCGGGCTGCTGAAGCCCACGGCCTGCACCATGCTGGTCAGCTCCCTCCGGGACCGCTT 674

Qy       2400 ATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTA 2459
Db         675 CCCCGACCTCCCACTGCACATCCACACCCACGACACGTCAGGGGCAGGCGTGGCAGCCAT 734

Qy       2460 CTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGG 2519
Db         735 GCTGGCCTGTGCCCAGGCTGGAGCTGATGTGGTGGATGTGGCAGCTGATTCCATGTCTGG 794

Qy       2520 CACCACCTCCCAGCCATCCCTGTCTGCCATTG 2551
Db         795 GATGACTTCACAGCCCAGCATGGGGGGCCCTG 826
```

RESULT 7
AQ875053/c
LOCUS AQ875053 859 bp DNA linear GSS 08-NOV-1999
DEFINITION V120F6 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION AQ875053
VERSION AQ875053.1 GI:6287297
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 859)
AUTHORS Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A., desEtages,S.A., Cheung,K.-H., Sheehan,A., Symoniatis,D., Jansen,R., Umansky,L., Heidtman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R., Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
JOURNAL Unpublished (1999)
COMMENT Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTTCTTTCTTTGGAAGTAC
Class: transposon-tagged.

```
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Saccharomyces cerevisiae"
                     /mol_type="genomic DNA"
                     /strain="Y2278 - S288C background, cir(0) rho(0)"
                     /db_xref="taxon:4932"
                     /lab_host="E. coli"
                     /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
                     /note="Vector: pHSS6-Sal; A yeast genomic DNA library
                     without 2 micron or mitochondrial DNA  was prepared in
                     pHSS6-Sal; genomic DNA was size-fractionated  (DNA of
                     roughly 2-3 kb in length) prior to cloning. This library
                     was subsequently mutagenized with a mTn-3xHA/lacZ
                     minitransposon  containing lacZ, URA3, and tet resistance.
                     "
ORIGIN

  Query Match              6.2%;  Score 223.4;  DB 28;  Length 859;
  Best Local Similarity   59.9%;  Pred. No. 4.5e-41;
  Matches  407;  Conservative    0;  Mismatches  268;  Indels    4;  Gaps    2;

Qy        827 AACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAG 886
              ||  | ||  ||  || |   |||||||||||||| ||  |      |||| ||   | |
Db        801 AAAATTCTTGGACAAAGCAAAAGCATATTGAAGTTCATTTGTGGCCCGATAACCACGAAA 742

Qy        887 AAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCG 946
              |||||  || ||||  ||| | |||||  ||  |||||  |    ||||||| ||||||
Db        741 CCGTTGTTCATCTTTTCGAAAGAGACTGTTCGGTGCAGAGAAA-CACCAAAAGGTTGTCA 683

Qy        947 AAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAG 1006
              ||  | ||  |||||  ||  |||     |  |||  |  ||||   |||||    | |||||||
Db        682 AAGTGGCNCCAGCAAAGACTTTACNCCGTGAAGTCGGTGACGCCATTTTGACAGATGCAG 623

Qy       1007 TAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATG 1066
              | ||  ||   |      |||  |||      |||||  ||  |  |||||||||| |||
Db        622 TTAAATTGGCCAAAGAGTGTGGCTACAGAAATGCGGGTACTGCTGAATTCTTGGTTGATA 563

Qy       1067 AAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGA 1126
              |  |    |  |||   |||||  |||||  ||  |||  |  |||||  ||  ||  ||| | |
Db        562 ACCAAAATAGACACTATTTCATTGAAATTAATCCAAGAATCCAAGTGGAACATACCATCA 503

Qy       1127 CTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAA 1186
              | ||||||  ||||    |  ||  |  |||  |||  |||||  |    |  ||  |||||
Db        502 CAGAAGAAATTACCGGTATAGATATTGTGGCGGCTCAGATCCAAATTGCGGCAGGTGCCT 443

Qy       1187 CCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGT 1246
              |  |    |  ||||  ||   |||  ||  ||  |  |||  |   ||    ||  |  ||||
Db        442 CTCTACCCCAGCTGGGCCTATTCCAGGACAAAATTACGACTCGTGGCTTTGCCATTCAGT 383

Qy       1247 GCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCG 1306
              ||||  ||  ||||||||||||  ||    ||   |||||  ||||||||||  |  ||    |
Db        382 GCCGTATTACCACGGAAGACCCTGCTAAGAACTTCCAACCAGATACCGGTAGAATAGAAG 323

Qy       1307 CGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTG 1363
              ||||||  ||   ||||  ||    |||  |||  |  ||  ||  ||||     ||   |    ||
Db        322 TGTACCGTTCTGCAGGTGGTAATGGTGTTAGACTGGATGGTGGTAACGCCTATGCAGGAA 263

Qy       1364 GCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1423
              ||||  |  | ||  |  |||||  |||||||||  |||    |  |||   ||||||  ||
Db        262 CAATAATCTCACCTCATTACGACTCAATGCTGGTCAAATGCTCATGCTCCGGTTCCACCT 203

Qy       1424 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1483
              ||||  |  |  |          ||  ||  ||   |||||||  |   |||||
Db        202 ACGAAATCGTTCGTAGAAAAATGATTCGTGCATTAATCGAGTTCAGAATTAGAGGTGTCA 143
```

```
Qy       1484 CAACCAACATTGGTTTCTT 1502
              |||||||||   ||| |
Db        142 AGACCAACATTCCCTTCCT 124
```

RESULT 8
BF251052

```
LOCUS       BF251052                637 bp    mRNA    linear   EST 15-NOV-2001
DEFINITION  EST418309 Coccidioides immitis spherule cDNA library Coccidioides
            immitis cDNA clone CIAAB85 5' sequence, mRNA sequence.
ACCESSION   BF251052
VERSION     BF251052.1  GI:16931195
KEYWORDS    EST.
SOURCE      Coccidioides immitis
  ORGANISM  Coccidioides immitis
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Gardner,M.J. and Kirkland,T.
  TITLE     Generation of ESTs from Coccidioides immitis spherule cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Malcolm J. Gardner
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Coccidioides immitis"
                     /mol_type="mRNA"
                     /db_xref="taxon:5501"
                     /clone="CIAAB85"
                     /dev_stage="spherule"
                     /lab_host="SOLR"
                     /clone_lib="Coccidioides immitis spherule cDNA library"
                     /note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2:
                     Xho1"
ORIGIN

  Query Match              6.1%;  Score 220.2;  DB 10;  Length 637;
  Best Local Similarity   60.4%;  Pred. No. 2.2e-40;
  Matches  380;  Conservative    0;  Mismatches  248;  Indels    1;  Gaps    1;

Qy        685 TACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGCGTTTTGTTGCT 744
              | |||||||  |  |||||  |    |||  ||  ||  ||  |||||  ||||||  | ||
Db          7 TTCCCCATCATCATCAAGGCCGCCTTCGGCGGCGGTGGCCGCGGCATGCGTGTCGTCCGC 66

Qy        745 TCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTC 804
                  |  ||    ||  |||  |  |     ||  |    ||  ||   ||  ||  |||
Db         67 GACCAGGAGTCCCTCCGCGACTCCTTCGAGCGCGCCACCTCGGAGGCCAAATCCGCCTTC 126

Qy        805 GGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAG 864
              |||  |  |||  |  ||  |  |||||  ||     |    |  ||   ||||  ||  ||  |||
Db        127 GGCAACGGCACCGTCTTCGTCGAGCGCTTCCTCGACAGGCCCAAGCACATCGAGGTCCAG 186

Qy        865 ATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAG 924
                |  ||  |||||   ||    ||   |  ||  ||  |||||  ||  ||  ||||||||||     | |||
Db        187 CTGCTCGGCGACAACCACGGCAACGTCGTCCACCTCTACGAGCGTGACTGCAGCGTCCAG 246

Qy        925 CGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGT 984
              ||  ||  |||||  ||  ||  |||||   |  ||  ||  ||   |   |     ||   ||.
Db        247 CGCCGCCACCAGAAGGTCGTCGAGCTCGCCCCGGCCAAGGACCTCCCCGTCGACGTCCGC 306

Qy        985 GATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGA 1044
              ||  ||||     ||  ||  ||  ||  |||  |   |   |||  |  |||||    |||  ||
Db        307 GACAGCATCCTCGCCGACGCCGTCAAGCTGGCCAAGTCCGTCAGCTACCGCAACGCCGGC 366

Qy       1045 ACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGT 1104
              ||||   ||  |||  |  |||||  |        |   |  ||  ||||||||||  ||  |||||  ||
Db        367 ACCGCCGAGTTCCTCGTCGACCAGCTCAACCGCTACTACTTCATCGAGATCAACCCCCGC 426
```

```
Qy       1105 ATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAG 1164
             |||||||| ||||||||| | || || || ||||||  | ||| | ||   || |||
Db        427 ATCCAGGTCGAGCACACCATCACCGAGGAGATCACCGGCATCGACATCGTCGCCGCCCAG 486

Qy       1165 ATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAG 1224
             || |  | || || || || ||| |  || |  | || || ||||| ||    |||
Db        487 ATCCAGATCGCCGCCGGCGCCACCCTCGAGCAGCTCGGCCTCACCCAGGACCGCATCTCC 546

Qy       1225 ACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGAT-CCAAACAACGGCTTCCG 1283
             |||   ||    ||  | ||||||||||||||||| |||||| ||| | |  |||||||
Db        547 ACCAGAGGCTTCGCCATCCAGTGCCGCATCACCACCGAAGATCCCACAAAGGGGCTTCCA 606

Qy       1284 CCCAGATACCGGAACTATCACCGCGTACC 1312
              || || |||||  |  |||   |  ||||
Db        607 GCCCGACACCGGTAAGATCGGAGTCTACC 635
```

RESULT 9
CA793824
LOCUS CA793824 862 bp mRNA linear EST 04-DEC-2002
DEFINITION AGENCOURT_11043532 NICHD_XGC_Emb1 Xenopus laevis cDNA clone IMAGE:6865891 5', mRNA sequence.
ACCESSION CA793824
VERSION CA793824.1 GI:26042921
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 862)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14488 row: p column: 18
High quality sequence stop: 649.
FEATURES Location/Qualifiers
source 1. .862
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6865891"
/tissue_type="embryo (stage 10)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_Emb1"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.55 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

Query Match 6.0%; Score 216.6; DB 14; Length 862;
Best Local Similarity 56.7%; Pred. No. 1.8e-39;
Matches 421; Conservative 0; Mismatches 319; Indels 3; Gaps 1;

```
Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900
              || ||||  || ||||||||| ||||| || || ||| |   |||| | || |||||  |
Db         20 AAACCTCGCCACATTGAAGTTCAGATACTCGGTGATAAATATGGAAATGTGGTACATTTG 79

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960
              |||||| | ||||| ||  | ||| |  | ||||| || || || ||||| || |||||
Db         80 TATGAAAGAGACTGTTCCGTCCAGAGAAGGCACCAGAAGGTGGTTGAAATAGCCCCAGCT 139
```

```
Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020
                || |||| |  | ||  | ||  |  |        ||  | |||||  |
Db        140 GCACAGCTGGACCTTCAGCTCAGAGACAGACTGACCCATGACTCTGTAAAACTGGCAAAA 199

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080
                | |||||  ||  || || || |||||  |||||||| ||| || | |||||  ||
Db        200 CAGGTGGGTTATGAGAATGCCGGCACTGTGGAGTTCTTGGTGGATAAACATGGCAAGCAT 259

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140
                |||||| ||| | |||  |  |  |||| |||||||| || || || |||||| | ||
Db        260 TACTTCATAGAAGTCAACTCCAGACTCCAAGTTGAGCATACAGTAACGGAAGAAATTACA 319

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200
              || || ||||| ||  | || |||||  |   | || |  ||   || | |    || |
Db        320 GATGTTGACCTAGTCCATGCCCAGATCCATGTTGCAGAGGGAAAAAGCCTTCCTGACCTT 379

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260
              || || |  ||||| || |||   | | | |||   ||  | ||||||  |  |||| ||
Db        380 GGCCTCAAGCAAGAGAACATCCGTATCAATGGTTGTGCCATTCAGTGCAGAGTCACAACA 439

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320
              |||||  ||         || || |  || |||||  ||     ||     |  | |||      |
Db        440 GAAGACCCTTCACGTGGGTTTCAGCCTGATACTGGTCGGATTGAGGTTTTCCGGAGTGGA 499

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCA 1377
              |  |||    ||  | ||  ||  |||  |||    |||||    |  ||  |  |  ||  || ||
Db        500 GAAGGAATGGGAATCCGACTAGACAGTGCTTCAGCTTTTCAAGGAGCTGTCATTTCCCCA 559

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437
              ||  | |||||||  | |||||  ||   |  |      |   ||      |||   | ||  |||  |
Db        560 CATTATGACTCCCTTCTGGTCAAGGTTATTGCACACGGAAAAGACCATCAAGTTGCCGCC 619

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGT 1497
                |          | ||  |||||  ||||||     |     ||||||   |  |||  |||
Db        620 ACAAAAATGAGCAGAGCTTTGGCAGAGTTCCGGATTAGGGGTGTTAAAGACAATATTCCC 679

Qy       1498 TTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTC 1557
              |||  |||      |  ||    |  |  |  ||                 | |  |||     ||
Db        680 TTCCTGCAAAACGTCTTGAACAATAATCAGTTTCTCCATGGTACAGTAGACACTCAGTTT 739

Qy       1558 ATTGCCGATCACCCGCACCTCCT 1580
              ||||  ||  | ||  |||||  |
Db        740 ATTGATGAAAATCCAGACCTCTT 762
```

RESULT 10
AI109467
LOCUS AI109467 675 bp mRNA linear EST 19-APR-2001
DEFINITION GH08678.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH08678 5prime, mRNA sequence.
ACCESSION AI109467
VERSION AI109467.1 GI:3477791
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 675)
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.
TITLE BDGP/HHMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 86 row: G column: 6
High quality sequence stop: 588.
FEATURES Location/Qualifiers
source 1. .675

```
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="GH08678"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
ORIGIN

  Query Match             6.0%;  Score 216.4;  DB 9;  Length 675;
  Best Local Similarity   58.9%;  Pred. No. 1.8e-39;
  Matches  392;  Conservative    0;  Mismatches  271;  Indels    3;  Gaps    1;

Qy       1719 ACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGA 1778
Db          8 ACCACCGAAGGGACTCCGCGAGGTCCTGGTGTGCGAAGGTCCGGAGGCCTTCGCCAAGGA 67

Qy       1779 TCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTC 1838
Db         68 GGTGCGCAACCGTAAGGAACTGCTACTCATGGACACCACCTTCCGCGATGCCCACCAGTC 127

Qy       1839 TTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAA 1898
Db        128 GCTGCTGGCCACCCGCGTGCGTTCCCACGATCTGCTGAAGATCTCCCCCTACGTGACGCA 187

Qy       1899 GCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGAT 1958
Db        188 CAAGTTCAACAACCTGTATTCGCTGGAGAACTGGGGCGGAGCCACCTTCGACGTGGCGCT 247

Qy       1959 GCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAA 2018
Db        248 GCGCTTCCTGCACGAGTGCCCGTGGGAGCGCCTGGAGGAGATGCGCAAGCGCATCCCGAA 307

Qy       2019 TGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACCCCGTACCCAGA 2078
Db        308 CATTCCCTTCCAGATGCTGTTGCGCGGAGCCAACGCCGTCGGCTACACCAGCTATCCGGA 367

Qy       2079 CTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCAT 2138
Db        368 CAACGTGGTCTACAAGTTCTGTGAGCTGGCTGTGCAGACCGGCATGGACATCTTCAGGGT 427

Qy       2139 CTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGAC 2198
Db        428 GTTCGACTCGCTCAACTACCTGCCCAACCTGATCCTCGGCATGGAAGCCGCTGGAAAGGC 487

Qy       2199 CAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATCCAAATGA 2258
Db        488 GGGCGGCGTGGT---GGAGGCTGCCATCTCCTATACCGGAGACGTCAGCGATCCCAAGCG 544

Qy       2259 AAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGC 2318
Db        545 CACCAAGTATGATCTGAAATACTACACTAACCTTGCCGATGAGCTGGTCAAGGCGGGCAC 604

Qy       2319 TCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCT 2378
Db        605 CCACGTGCTCTGCATAAAGGATATGGCTGGTCTGCTCAAGCCCGAATCCGCCAGGCTTCT 664

Qy       2379 GGTCAC 2384
Db        665 GATCAC 670


RESULT 11
BI103175
LOCUS       BI103175                744 bp    mRNA    linear   EST 26-JUN-2001
DEFINITION  602889415F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044779
            5', mRNA sequence.
ACCESSION   BI103175
VERSION     BI103175.1  GI:14554068
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11122  row: m  column: 04
            High quality sequence stop: 737.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5044779"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Kid14"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
ORIGIN

  Query Match             5.9%;  Score 214.8;  DB 12;  Length 744;
  Best Local Similarity   58.6%;  Pred. No. 4.4e-39;
  Matches  430;  Conservative    0;  Mismatches  297;  Indels    7;  Gaps    3;

Qy       1765 GCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGC 1824
Db          15 GGGCTTGCCCGAGCTGTGCGGAATCACCACGGGCTGCTGTTGATGGACACAACCTTCCGG 74

Qy       1825 GATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCA 1884
Db          75 GATGCCCACCAGTCACTACTGGCCACTAGAGTGCGCACACATGATCTCAA-AAAGATGCG 133

Qy       1885 GAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACC 1944
Db         134 CCCTATGTTGCCCACAACTTCAACAAGCTCTTCAGCATGGAGAACTGGGGAGGCGCCACG 193

Qy       1945 TACGATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGC 2004
Db         194 TTCGACGTTGCCATGCGCTTCCTGTACGAGTGCCCCTGGCGGCGGCTCCAGGAGCTCCGG 253

Qy       2005 GAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATAC 2064
Db         254 GAGCTTATCCCGAACATCCCGTTCCAGATGCTACTGAGGGGGGCCAATGCTGTGGGCTAC 313

Qy       2065 ACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTG 2124
Db         314 ACCAACTACCCTGACAACGTGGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGAATGGTATG 373

Qy       2125 GACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGAC 2184
Db         374 GACGTCTTCCGAGTCTTTGACTCCCTCAACTACTTGCCAAACATGCTGCTGGGCATGGAA 433

Qy       2185 GCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTC 2244
Db         434 GCAG---CAGGCAGTGCTGGGGGTGTGGTGGAGGCTGCCATCTCATACACGGGGGACGTG 490

Qy       2245 TCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATC 2304
```

```
Db        491 GCTGACCCTAGTCGCACTAAATACTCACTGGAGTACTACATGGGCTTAGCTGAAGAACTG 550

Qy       2305 GTCAAGTCTGGCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCT 2364

Db        551 GTGCGAGCTGGCACTCACATCCTGTGCATTAAGGACATGGCGGGCCTGCTGAAGCCTGCC 610

Qy       2365 GCGGTAACCAAGCTGGTCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTG 2421

Db        611 GCCTGCACCATGCTGGTCAGCTCCCTCCGGGACCGATTCCCCGACCTCCCACTGCACATC 670

Qy       2422 CACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGT 2481

Db        671 CATACCCATGATACATCAGGGGCAGGTGTGGCAGCCATGCTGGCCTGTGCACAAGCAGGG 730

Qy       2482 GCAGATGCTGTTGA 2495

Db        731 GCTGATGTCGTGGA 744
```

RESULT 12
CK000115
LOCUS CK000115 800 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT_16368972 NIH_MGC_220 Homo sapiens cDNA clone IMAGE:30707731 5', mRNA sequence.
ACCESSION CK000115
VERSION CK000115.1 GI:38526149
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM1072 row: m column: 20
High quality sequence start: 12
High quality sequence stop: 654.
FEATURES Location/Qualifiers
source 1. .800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30707731"
/lab_host="DH10B TonA"
/clone_lib="NIH_MGC_220"
/note="Organ: mixed; Vector: pYX-Asc; Site_1: EcoRI; Site_2: NotI; Library is oligo-dT primed and directionally cloned Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according tomRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 0.5-1Kb. Adaptors 5'(AATTCGGCACGAGG)3' and 5'd (CCTCGTGCCG)3'. 3' Linker sequence - GCGGCCGCTGAGAGCC T18. Sequencing primers 3'end: T3 promoter primer 5'd (ATTAACCCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd (TAATACGACTCACTATAGGG)3'. Library was constructed in the laboratory of M. Bento Soares. Average insert size 3-4kb Note: this is a NIH_MGC Library."

ORIGIN

Query Match 5.9%; Score 213.2; DB 14; Length 800;
Best Local Similarity 59.1%; Pred. No. 1.1e-38;
Matches 401; Conservative 0; Mismatches 273; Indels 4; Gaps 2;

```
Qy        841 AACCCTCAGCATATTGAAGTGCAG-ATCCTTGGCGATCACACTGGAGAAGTTGTACACCT 899

Db         35 AAGCCACGGCACATCGAGGTGCAGAATCTTGGGGGACCAGTATGGGAACATCCTGCACCT 94

Qy        900 TTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGC 959

Db         95 GTACGAGCGAGACTGCTCCATCCAGCGGCGGCACCAGAAGGTGGTCGAGATTGCCCCCGC 154

Qy        960 ACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCG 1019

Db        155 CGCCCACCTGGACCCGCAGCTTCGGACTCGGCTCACCAGCGACTCTGTGAAACTCGCTAA 214

Qy       1020 CTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCA 1079

Db        215 ACAGGTGGGCTACGAGAACGCAGGCACCGTGGAGTTCCTGGTGGACAGGCACGGCAAGCA 274

Qy       1080 CGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCAC 1139

Db        275 CTACTTCATCGAGGTCAACTCCCGCCTGCAGGTGGAGCACACGGTCACAGAGGAGATCAC 334

Qy       1140 CGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATT 1199

Db        335 CGACGTAGACCTGGTCCATGCTCAGATCCACGTGGCTGAGGGCAGGAGCCTACCCGACCT 394

Qy       1200 GGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCAC 1259

Db        395 GGGCCTGCGGCAGGAGAACATCCGCATCAACGGGTGTGCCATCCAGTGCCGGGTCACCAC 454

Qy       1260 GGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACC 1319

Db        455 CGAGGACCCCGCGCGCAGCTTCCAGCCGGACACCGGCCGCATTGAGGTGTTCCGGAGCGG 514

Qy       1320 AGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGC 1376

Db        515 AGAGGGCATGGGCATCCGCCTGGATAATGCTTCCGCCTTCCAAGGAGCCGTCATCTCGCC 574

Qy       1377 ACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGT 1436

Db        575 CCACTACGACTCCCTGCTGGTCAAAGTCATTGCCCACGGCAAAGACCACCCCACGGCCGC 634

Qy       1437 TGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGG 1496

Db        635 CACCAAGATGAGCAGGGCCCTTGCGGAGTTCCGCGTCCGAGGTGTGAAGACCAACATCGC 694

Qy       1497 TTTCTTGCGTGCGTTGCT 1514

Db        695 CTTCCTGCAGAATGTGCT 712
```

RESULT 13
BI618593
LOCUS BI618593 640 bp mRNA linear EST 07-SEP-2001
DEFINITION RH49459.5prime RH Drosophila melanogaster normalized Head pFlc-1 Drosophila melanogaster cDNA clone RH49459 5 similar to BcDNA:GH06348: FBan0001516 GO:[pyruvate carboxylase (GO:0004736)] located on: 2R 46B14-46C1;: 08/19/2001, mRNA sequence.
ACCESSION BI618593
VERSION BI618593.1 GI:15514118
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 640)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,

```
            George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,
            Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and
            Rubin,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003831: arm:2R [4876906,5137815]
            estimated-cyto:46B13-46E4: 08/19/2001
            Plate: RH.494  row: E  column: 11
            High quality sequence stop: 543.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="RH49459"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
ORIGIN

  Query Match             5.8%;  Score 211;  DB 12;  Length 640;
  Best Local Similarity   59.4%;  Pred. No. 3.2e-38;
  Matches  377;  Conservative    0;  Mismatches  255;  Indels    3;  Gaps    1;

Qy        790 GCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAG 849
              ||  | ||  ||  ||  |||  ||||  |||  | |    |  ||          |        ||  |
Db          1 GCCAAGGCCGCCTTTGGCAATGGTGCGATGTTCATTGAGAAGTTCATCGAGCGACCGCGT 60

Qy        850 CATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGT 909
              ||  |||||  ||  ||   |  ||  ||  ||   |   | ||   | ||  ||  ||  ||  ||  ||||||
Db         61 CACATTGAGGTTCAACTGCTGGGAGACAAGGCCGGGAATGTAGTGCATCTGTACGAACGT 120

Qy        910 GACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTG 969
              ||||||||   ||||||||||  |||||  ||  ||  ||  ||  ||  |||||  |   |    ||
Db        121 GACTGCTCCGTGCAGCGTCGCCACCAGAAGGTGGTGGAGATCGCCCCAGCGCCACGCCTG 180

Qy        970 GATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGT 1029
                  |||  |  ||  ||     ||     |  ||  |||||     ||    ||||     |  ||
Db        181 CCCATCGAAATCCGGGACAAGATGACGGAGGCAGCAGTGCGCTTGGCCCGCCATGTGGGC 240

Qy       1030 TACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATC 1089
              ||  |   |||  ||||||||||||  |||  |   |||  ||    ||||||   |  |||||||
Db        241 TATGAAAACGCCGGAACCGTGGAGTTCCTTTGCGACGAGTCCGGCAACTTCTACTTCATC 300

Qy       1090 GAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGAC 1149
              |||  |||||  |  |||  |  ||  ||  ||||||||  ||  ||  ||  ||  |  ||   |   |  ||
Db        301 GAAGTGAACGCCCGTCTACAAGTGGAGCACACGGTCACCGAGGAGATAACAGGCATCGAT 360

Qy       1150 CTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACC 1209
                |||||  ||  ||||  ||  |||  ||||  |   ||     |||  |    ||||||||     ||
Db        361 TTGGTGCAGTCGCAAATCCGCGTGGCCGAGGGCATGACCCTACCCGAATTGGGCTACACG 420

Qy       1210 CAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCA 1269
              ||  ||  ||||||   |  |  |  ||    ||  |  ||||||||   |  |||||  ||||||||
Db        421 CAGGACAAGATCGTGCCGCGTGGCTATGCCATTCAGTGCCGTGTGACCACCGAAGATCCG 480

Qy       1270 AACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCT 1329
                |||||   ||  |   ||  |  |||||      |     |  |  |||  ||     |  ||
Db        481 GCCAACGATTTTCAGCCCAACACCGGTCGCTTGGAGGTCTTCCGATCTGGCGAGGGTATG 540

Qy       1330 GGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGAC 1386
              |||  ||  |   |  |||  |||    |  ||   |    |||  |  |    ||   |   ||  |||   |||
Db        541 GGCATTAGGTTGGACAGTGCGTCCGCCTACGCCGGAGCCATCATTTCGCCATACTACGAC 600

Qy       1387 TCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGA 1421
              |||  |||||||  ||   |||  ||   |   |    |||
Db        601 TCCCTGCTGGTCAAGGTGATCTCGCACGCCAGCGA 635


RESULT 14
BG910468
LOCUS       BG910468                715 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602806142F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4938203
            5', mRNA sequence.
ACCESSION   BG910468
VERSION     BG910468.1  GI:14290944
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10873  row: d  column: 12
            High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4938203"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn67"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             5.7%;  Score 208.2;  DB 12;  Length 715;
  Best Local Similarity   56.8%;  Pred. No. 1.5e-37;
  Matches  403;  Conservative    0;  Mismatches  303;  Indels    3;  Gaps    1;

Qy        872 GCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTC 931
              | ||  ||   |||  |  |  |||||  ||  ||  ||||||||  |  ||||| ||  |
Db          1 GGGACCAGTATGGGAACATCCTGCACCTGTACGAGCGAGACTGCTCCATCCAGCGGCGGC 60

Qy        932 ACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCA 991
              ||||  ||  ||  |||||  |||||  ||  ||    ||  ||||  ||  |  ||  ||    |||
Db         61 ACCAGAAGGTGGTCGAGATTGCCCCCGCCGCCCACCTGGACCCGCAGCTTCGGACTCGGC 120

Qy        992 TTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGG 1051
              |         ||  |  |||  ||  ||          |  ||  |||  ||  |||  ||  |||||||
Db        121 TCACCAGCGACTCTGTGAAACTCGCTAAACAGGTGGGCTACGAGAACGCAGGCACCGTGG 180

Qy       1052 AATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGG 1111
              |  |||  ||||  ||      |  |||||  |||  |||||||||   |  |||  |  ||   |  ||||
Db        181 AGTTCCTGGTGGACAGGCACGGCAAGCACTACTTCATCGAGGTCAACTCCCGCCTGCAGG 240

Qy       1112 TTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCT 1171
```

```
Db        241 TGGAGCACACGGTCACAGAGGAGATCACCGACGTAGACCTGGTCCATGCTCAGATCCACG 300

Qy       1172 TGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACG 1231

Db        301 TGGCTGAGGGCAGGAGCCTACCCGACCTGGGCCTGCGGCAGGAGAACATCCGCATCAACG 360

Qy       1232 GTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATA 1291

Db        361 GGTGTGCCATCCAGTGCCGGGTCACCACCGAGGACCCCGCGCGCAGCTTCCAGCCGGACA 420

Qy       1292 CCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAG 1351

Db        421 CCGGCCGCATTGAGGTGTTCCGGAGCGGAGAGGGCATGGGCATCCGCCTGGATAATGCTT 480

Qy       1352 CTCAGCTC---GGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCT 1408

Db        481 CCGCCTTCCAAGGAGCCGTCATCTCGCCCCACTACGACTCCCTGCTGGTCAAAGTCATTG 540

Qy       1409 GCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCA 1468

Db        541 CCCACGGCAAAGACCAACCCACGGCCGCCACCAAGATGAGCAGGGCCCTTGCGGAGTTCC 600

Qy       1469 CCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGACT 1528

Db        601 GCGTCCGAGGTGTGAAGACCAACATCGCCTTCCTGCAGAATGTGCTCAACAACCAGCAGT 660

Qy       1529 TCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCT 1577

Db        661 TCCTGGCAGGCACTGTGGACACCAGTTCATCGACGCAGAACCCAGAGCT 709
```

```
RESULT 15
BJ063722
LOCUS       BJ063722                641 bp    mRNA    linear   EST 29-SEP-2003
DEFINITION  BJ063722 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL077c10 5', mRNA sequence.
ACCESSION   BJ063722
VERSION     BJ063722.1  GI:17470882
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL077c10"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
ORIGIN

  Query Match             5.6%;  Score 204.4;  DB 12;  Length 641;
```

```
  Best Local Similarity   58.9%;  Pred. No. 1.1e-36;
  Matches  371;  Conservative    0;  Mismatches  256;  Indels    3;  Gaps    1;

Qy        841 AACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGTTGTACACCTT 900

Db          5 AAACCTCGCCACATTGAAGTTCAGATACTCGGTGATAAATATGGAAATGTGGTACATTTG 64

Qy        901 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCA 960

Db         65 TATGAAAGAGACTGTTCCGTCCAGAGAAGGCACCAGAAGGTGGTTGAAATAGCCCCAGCT 124

Qy        961 CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGC 1020

Db        125 GCACAGCTGGACCTTCAGCTCAGAGACAGACTGACCCATGACTCTGTAAAACTGGCAAAA 184

Qy       1021 TCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC 1080

Db        185 CAGGTGGGTTATGAGAATGCCGGCACTGTGGAGTTCTTGGTGGATAAACATGGCAAGCAT 244

Qy       1081 GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC 1140

Db        245 TACTTCATAGAAGTCAACTCCAGACTCCAAGTTGAGCATACAGTAACGGAAGAAATTACA 304

Qy       1141 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTG 1200

Db        305 GATGTTGACCTAGTCCATGCCCAGATCCATGTTGCAGAGGGAAAAAGCCTTCCTGACCTT 364

Qy       1201 GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCCGCATCACCACG 1260

Db        365 GGCCTCAAGCAAGAGAACATCCGTATCAATGGTTGTGCCATTCAGTGCAGAGTCACAACA 424

Qy       1261 GAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA 1320

Db        425 GAAGACCCTTCACGTGGGTTTCAGCCTGATACTGGTCGGATTGAGGTTTTCCGGAGTGGA 484

Qy       1321 GGCGGAGCTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCA 1377

Db        485 GAAGGAATGGGAATCCGACTAGACAGTGCTTCAGCTTTTCAAGGAGCTGTCATTTCCCCA 544

Qy       1378 CACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1437

Db        545 CATTATGACTCCCTTCTGGTCAAGGTTATTGCACACGGAAAAGACCATCAAGTTGCCGCC 604

Qy       1438 GCTCGTGCACAGCGCGCGTTGGCTGAGTTC 1467

Db        605 ACAAAAATGAGCAGAGCTTTGGCAGAGTTC 634
```

```
Search completed: March 24, 2004, 02:37:29
Job time : 8678 secs
```